University of Co-operative and Management,
Department of Statistics

Department of Statistics

INTRODUCTION TO RESEARCH METHODOLOGY

MAR 111

RESEARCH METHODOLOGY

FIRST, QUARTER

UNIVERSITY OF CO-OPERATIVE AND MANAGEMENT, THANLYIN
DEPARTMENT OF STATISTICS

INTRODUCTOIN TO RESEARCH METHODOLOGY
MAR 111

FIRST, QUARTER

Contents

Preface to the Second Edition *vii*
Preface to the First Edition *ix*

1. Research Methodology: An Introduction 1
   - Meaning of Research *1*
   - Objectives of Research *2*
   - Motivation in Research *2*
   - Types of Research *2*
   - Research Approaches *5*
   - Significance of Research *5*
   - Research Methods versus Methodology *7*
   - Research and Scientific Method *9*
   - Importance of Knowing How Research is Done *10*
   - Research Process *10*
   - Criteria of Good Research *20*
   - Problems Encountered by Researchers in India *21*

2. Defining the Research Problem 24
   - What is a Research Problem? *24*
   - Selecting the Problem *25*
   - Necessity of Defining the Problem *26*
   - Technique Involved in Defining a Problem *27*
   - An Illustration *29*
   - Conclusion *29*

3. Research Design 31
   - Meaning of Research Design *31*
   - Need for Research Design *32*

Features of a Good Design *33*
Important Concepts Relating to Research Design *33*
Different Research Designs *35*
Basic Principles of Experimental Designs *39*
Conclusion *52*
Appendix
Developing a Research Plan *53*

4. Sampling Design 55
Census and Sample Survey *55*
Implications of a Sample Design *55*
Steps in Sampling Design *56*
Criteria of Selecting a Sampling Procedure *57*
Characteristics of a Good Sample Design *58*
Different Types of Sample Designs *58*
How to Select a Random Sample? *60*
Random Sample from an Infinite Universe *61*
Complex Random Sampling Designs *62*
Conclusion *67*

5. Measurement and Scaling Techniques 69
Measurement in Research *69*
Measurement Scales *71*
Sources of Error in Measurement *72*
Tests of Sound Measurement *73*
Technique of Developing Measurement Tools *75*
Scaling *76*
Meaning of Scaling *76*
Scale Classification Bases *77*
Important Scaling Techniques *78*
Scale Construction Techniques *82*

6. Methods of Data Collection 95
Collection of Primary Data *95*
Observation Method *96*
Interview Method *97*
Collection of Data through Questionnaires *100*
Collection of Data through Schedules *104*
Difference between Questionnaires and Schedules *104*
Some Other Methods of Data Collection *106*
Collection of Secondary Data *111*

Selection of Appropriate Method for Data Collection *112*
Case Study Method *113*
Appendices
(i) Guidelines for Constructing Questionnaire/Schedule *118*
(ii) Guidelines for Successful Interviewing *119*
(iii) Difference between Survey and Experiment *120*

1

Research Methodology: An Introduction

MEANING OF RESEARCH

Research in common parlance refers to a search for knowledge. Once can also define research as a scientific and systematic search for pertinent information on a specific topic. In fact, research is an art of scientific investigation. The Advanced Learner's Dictionary of Current English lays down the meaning of research as "a careful investigation or inquiry specially through search for new facts in any branch of knowledge."[1] Redman and Mory define research as a "systematized effort to gain new knowledge."[2] Some people consider research as a movement, a movement from the known to the unknown. It is actually a voyage of discovery. We all possess the vital instinct of inquisitiveness for, when the unknown confronts us, we wonder and our inquisitiveness makes us probe and attain full and fuller understanding of the unknown. This inquisitiveness is the mother of all knowledge and the method, which man employs for obtaining the knowledge of whatever the unknown, can be termed as research.

Research is an academic activity and as such the term should be used in a technical sense. According to Clifford Woody research comprises defining and redefining problems, formulating hypothesis or suggested solutions; collecting, organising and evaluating data; making deductions and reaching conclusions; and at last carefully testing the conclusions to determine whether they fit the formulating hypothesis. D. Slesinger and M. Stephenson in the Encyclopaedia of Social Sciences define research as "the manipulation of things, concepts or symbols for the purpose of generalising to extend, correct or verify knowledge, whether that knowledge aids in construction of theory or in the practice of an art."[3] Research is, thus, an original contribution to the existing stock of knowledge making for its advancement. It is the persuit of truth with the help of study, observation, comparison and experiment. In short, the search for knowledge through objective and systematic method of finding solution to a problem is research. The systematic approach concerning generalisation and the formulation of a theory is also research. As such the term 'research' refers to the systematic method

[1] *The Advanced Learner's Dictionary of Current English*, Oxford, 1952, p. 1069.

[2] L.V. Redman and A.V.H. Mory, *The Romance of Research*, 1923, p.10.

[3] *The Encyclopaedia of Social Sciences*, Vol. IX, MacMillan, 1930.

consisting of enunciating the problem, formulating a hypothesis, collecting the facts or data, analysing the facts and reaching certain conclusions either in the form of solutions(s) towards the concerned problem or in certain generalisations for some theoretical formulation.

OBJECTIVES OF RESEARCH

The purpose of research is to discover answers to questions through the application of scientific procedures. The main aim of research is to find out the truth which is hidden and which has not been discovered as yet. Though each research study has its own specific purpose, we may think of research objectives as falling into a number of following broad groupings:

1. To gain familiarity with a phenomenon or to achieve new insights into it (studies with this object in view are termed as *exploratory* or *formulative* research studies);
2. To portray accurately the characteristics of a particular individual, situation or a group (studies with this object in view are known as *descriptive* research studies);
3. To determine the frequency with which something occurs or with which it is associated with something else (studies with this object in view are known as *diagnostic* research studies);
4. To test a hypothesis of a causal relationship between variables (such studies are known as *hypothesis-testing* research studies).

MOTIVATION IN RESEARCH

What makes people to undertake research? This is a question of fundamental importance. The possible motives for doing research may be either one or more of the following:

1. Desire to get a research degree along with its consequential benefits;
2. Desire to face the challenge in solving the unsolved problems, i.e., concern over practical problems initiates research;
3. Desire to get intellectual joy of doing some creative work;
4. Desire to be of service to society;
5. Desire to get respectability.

However, this is not an exhaustive list of factors motivating people to undertake research studies. Many more factors such as directives of government, employment conditions, curiosity about new things, desire to understand causal relationships, social thinking and awakening, and the like may as well motivate (or at times compel) people to perform research operations.

TYPES OF RESEARCH

The basic types of research are as follows:

(i) *Descriptive vs. Analytical: Descriptive research* includes surveys and fact-finding enquiries of different kinds. The major purpose of descriptive research is description of the state of affairs as it exists at present. In social science and business research we quite often use

the term *Ex post facto research* for descriptive research studies. The main characteristic of this method is that the researcher has no control over the variables; he can only report what has happened or what is happening. Most *ex post facto research* projects are used for descriptive studies in which the researcher seeks to measure such items as, for example, frequency of shopping, preferences of people, or similar data. *Ex post facto studies* also include attempts by researchers to discover causes even when they cannot control the variables. The methods of research utilized in descriptive research are survey methods of all kinds, including comparative and correlational methods. In *analytical research*, on the other hand, the researcher has to use facts or information already available, and analyze these to make a critical evaluation of the material.

(ii) *Applied vs. Fundamental:* Research can either be applied (or action) research or fundamental (to basic or pure) research. *Applied research* aims at finding a solution for an immediate problem facing a society or an industrial/business organisation, whereas *fundamental research* is mainly concerned with generalisations and with the formulation of a theory. "Gathering knowledge for knowledge's sake is termed 'pure' or 'basic' research."[4] Research concerning some natural phenomenon or relating to pure mathematics are examples of fundamental research. Similarly, research studies, concerning human behaviour carried on with a view to make generalisations about human behaviour, are also examples of fundamental research, but research aimed at certain conclusions (say, a solution) facing a concrete social or business problem is an example of applied research. Research to identify social, economic or political trends that may affect a particular institution or the copy research (research to find out whether certain communications will be read and understood) or the marketing research or evaluation research are examples of applied research. Thus, the central aim of applied research is to discover a solution for some pressing practical problem, whereas basic research is directed towards finding information that has a broad base of applications and thus, adds to the already existing organized body of scientific knowledge.

(iii) *Quantitative vs. Qualitative:* Quantitative research is based on the measurement of quantity or amount. It is applicable to phenomena that can be expressed in terms of quantity. Qualitative research, on the other hand, is concerned with qualitative phenomenon, i.e., phenomena relating to or involving quality or kind. For instance, when we are interested in investigating the reasons for human behaviour (i.e., why people think or do certain things), we quite often talk of 'Motivation Research', an important type of qualitative research. This type of research aims at discovering the underlying motives and desires, using in depth interviews for the purpose. Other techniques of such research are word association tests, sentence completion tests, story completion tests and similar other projective techniques. Attitude or opinion research i.e., research designed to find out how people feel or what they think about a particular subject or institution is also qualitative research. Qualitative research is specially important in the behavioural sciences where the aim is to discover the underlying motives of human behaviour. Through such research we can analyse the various factors which motivate people to behave in a particular manner or which make people like or dislike a particular thing. It may be stated, however, that to apply qualitative research in

[4] Pauline V. Young, *Scientific Social Surveys and Research*, p. 30.

practice is relatively a difficult job and therefore, while doing such research, one should seek guidance from experimental psychologists.

(iv) *Conceptual vs. Empirical:* Conceptual research is that related to some abstract idea(s) or theory. It is generally used by philosophers and thinkers to develop new concepts or to reinterpret existing ones. On the other hand, empirical research relies on experience or observation alone, often without due regard for system and theory. It is data-based research, coming up with conclusions which are capable of being verified by observation or experiment. We can also call it as experimental type of research. In such a research it is necessary to get at facts firsthand, at their source, and actively to go about doing certain things to stimulate the production of desired information. In such a research, the researcher must first provide himself with a working hypothesis or guess as to the probable results. He then works to get enough facts (data) to prove or disprove his hypothesis. He then sets up experimental designs which he thinks will manipulate the persons or the materials concerned so as to bring forth the desired information. Such research is thus characterised by the experimenter's control over the variables under study and his deliberate manipulation of one of them to study its effects. Empirical research is appropriate when proof is sought that certain variables affect other variables in some way. Evidence gathered through experiments or empirical studies is today considered to be the most powerful support possible for a given hypothesis.

(v) *Some Other Types of Research:* All other types of research are variations of one or more of the above stated approaches, based on either the purpose of research, or the time required to accomplish research, on the environment in which research is done, or on the basis of some other similar factor. Form the point of view of time, we can think of research either as *one-time research or longitudinal research.* In the former case the research is confined to a single time-period, whereas in the latter case the research is carried on over several time-periods. Research can be *field-setting research or laboratory research or simulation research*, depending upon the environment in which it is to be carried out. Research can as well be understood as *clinical or diagnostic research.* Such research follow case-study methods or indepth approaches to reach the basic causal relations. Such studies usually go deep into the causes of things or events that interest us, using very small samples and very deep probing data gathering devices. The research may be *exploratory* or it may be formalized. The objective of exploratory research is the development of hypotheses rather than their testing, whereas formalized research studies are those with substantial structure and with specific hypotheses to be tested. *Historical research* is that which utilizes historical sources like documents, remains, etc. to study events or ideas of the past, including the philosophy of persons and groups at any remote point of time. Research can also be classified as *conclusion-oriented* and decision-oriented. While doing conclusion-oriented research, a researcher is free to pick up a problem, redesign the enquiry as he proceeds and is prepared to conceptualize as he wishes. Decision-oriented research is always for the need of a decision maker and the researcher in this case is not free to embark upon research according to his own inclination. Operations research is an example of decision oriented research since it is a scientific method of providing executive departments with a quantitative basis for decisions regarding operations under their control.

Research Approaches

The above description of the types of research brings to light the fact that there are two basic approaches to research, viz., *quantitative approach* and the *qualitative approach*. The former involves the generation of data in quantitative form which can be subjected to rigorous quantitative analysis in a formal and rigid fashion. This approach can be further sub-classified into *inferential*, *experimental* and *simulation approaches* to research. The purpose of *inferential approach* to research is to form a data base from which to infer characteristics or relationships of population. This usually means survey research where a sample of population is studied (questioned or observed) to determine its characteristics, and it is then inferred that the population has the same characteristics. *Experimental approach* is characterised by much greater control over the research environment and in this case some variables are manipulated to observe their effect on other variables. *Simulation approach* involves the construction of an artificial environment within which relevant information and data can be generated. This permits an observation of the dynamic behaviour of a system (or its sub-system) under controlled conditions. The term 'simulation' in the context of business and social sciences applications refers to "the operation of a numerical model that represents the structure of a dynamic process. Given the values of initial conditions, parameters and exogenous variables, a simulation is run to represent the behaviour of the process over time."[5] Simulation approach can also be useful in building models for understanding future conditions.

Qualitative approach to research is concerned with subjective assessment of attitudes, opinions and behaviour. Research in such a situation is a function of researcher's insights and impressions. Such an approach to research generates results either in non-quantitative form or in the form which are not subjected to rigorous quantitative analysis. Generally, the techniques of focus group interviews, projective techniques and depth interviews are used. All these are explained at length in chapters that follow.

Significance of Research

"All progress is born of inquiry. Doubt is often better than overconfidence, for it leads to inquiry, and inquiry leads to invention" is a famous Hudson Maxim in context of which the significance of research can well be understood. Increased amounts of research make progress possible. *Research inculcates scientific and inductive thinking and it promotes the development of logical habits of thinking and organisation.*

The role of research in several fields of applied economics, whether related to business or to the economy as a whole, has greatly increased in modern times. The increasingly complex nature of business and government has focused attention on the use of research in solving operational problems. Research, as an aid to economic policy, has gained added importance, both for government and business.

Research provides the basis for nearly all government policies in our economic system. For instance, government's budgets rest in part on an analysis of the needs and desires of the people and on the availability of revenues to meet these needs. The cost of needs has to be equated to probable revenues and this is a field where research is most needed. Through research we can devise alternative policies and can as well examine the consequences of each of these alternatives.

[5] Robert C. Meir, William T. Newell and Harold L. Dazier, *Simulation in Business and Economics*, p. 1.

Decision-making may not be a part of research, but research certainly facilitates the decisions of the policy maker. Government has also to chalk out programmes for dealing with all facets of the country's existence and most of these will be related directly or indirectly to economic conditions. The plight of cultivators, the problems of big and small business and industry, working conditions, trade union activities, the problems of distribution, even the size and nature of defence services are matters requiring research. Thus, research is considered necessary with regard to the allocation of nation's resources. Another area in government, where research is necessary, is collecting information on the economic and social structure of the nation. Such information indicates what is happening in the economy and what changes are taking place. Collecting such statistical information is by no means a routine task, but it involves a variety of research problems. These day nearly all governments maintain large staff of research technicians or experts to carry on this work. Thus, in the context of government, research as a tool to economic policy has three distinct phases of operation, viz., (i) investigation of economic structure through continual compilation of facts; (ii) diagnosis of events that are taking place and the analysis of the forces underlying them; and (iii) the prognosis, i.e., the prediction of future developments.

Research has its special significance in solving various operational and planning problems of business and industry. Operations research and market research, along with motivational research, are considered crucial and their results assist, in more than one way, in taking business decisions. Market research is the investigation of the structure and development of a market for the purpose of formulating efficient policies for purchasing, production and sales. Operations research refers to the application of mathematical, logical and analytical techniques to the solution of business problems of cost minimisation or of profit maximisation or what can be termed as optimisation problems. Motivational research of determining why people behave as they do is mainly concerned with market characteristics. In other words, it is concerned with the determination of motivations underlying the consumer (market) behaviour. All these are of great help to people in business and industry who are responsible for taking business decisions. Research with regard to demand and market factors has great utility in business. Given knowledge of future demand, it is generally not difficult for a firm, or for an industry to adjust its supply schedule within the limits of its projected capacity. Market analysis has become an integral tool of business policy these days. Business budgeting, which ultimately results in a projected profit and loss account, is based mainly on sales estimates which in turn depends on business research. Once sales forecasting is done, efficient production and investment programmes can be set up around which are grouped the purchasing and financing plans. Research, thus, replaces intuitive business decisions by more logical and scientific decisions.

Research is equally important for social scientists in studying social relationships and in seeking answers to various social problems. It provides the intellectual satisfaction of knowing a few things just for the sake of knowledge and also has practical utility for the social scientist to know for the sake of being able to do something better or in a more efficient manner. Research in social sciences is concerned both with knowledge for its own sake and with knowledge for what it can contribute to practical concerns. "This double emphasis is perhaps especially appropriate in the case of social science. On the one hand, its responsibility as a science is to develop a body of principles that make possible the understanding and prediction of the whole range of human interactions. On the other hand, because of its social orientation, it is increasingly being looked to for practical guidance in solving immediate problems of human relations."[6]

[6] Marie Jahoda, Morton Deutsch and Stuart W. Cook, *Research Methods in Social Relations*, p. 4.

In addition to what has been stated above, the significance of research can also be understood keeping in view the following points:

(a) To those students who are to write a master's or Ph.D. thesis, research may mean a careerism or a way to attain a high position in the social structure;
(b) To professionals in research methodology, research may mean a source of livelihood;
(c) To philosophers and thinkers, research may mean the outlet for new ideas and insights;
(d) To literary men and women, research may mean the development of new styles and creative work;
(e) To analysts and intellectuals, research may mean the generalisations of new theories.

Thus, research is the fountain of knowledge for the sake of knowledge and an important source of providing guidelines for solving different business, governmental and social problems. It is a sort of formal training which enables one to understand the new developments in one's field in a better way.

Research Methods versus Methodology

It seems appropriate at this juncture to explain the difference between research methods and research methodology. *Research methods* may be understood as all those methods/techniques that are used for conduction of research. *Research methods or techniques*, thus, refer to the methods the researchers*

*At times, a distinction is also made between research techniques and research methods. *Research techniques* refer to the behaviour and instruments we use in performing research operations such as making observations, recording data, techniques of processing data and the like. *Research methods* refer to the behaviour and instruments used in selecting and constructing research technique. For instance, the difference between methods and techniques of data collection can better be understood from the details given in the following chart—

Type	*Methods*	*Techniques*
1. Library Research	(i) Analysis of historical records	Recording of notes, Content analysis, Tape and Film listening and analysis.
	(ii) Analysis of documents	Statistical compilations and manipulations, reference and abstract guides, contents analysis.
2. Field Research	(i) Non-participant direct observation	Observational behavioural scales, use of score cards, etc.
	(ii) Participant observation	Interactional recording, possible use of tape recorders, photo graphic techniques.
	(iii) Mass observation	Recording mass behaviour, interview using independent observers in public places.
	(iv) Mail questionnaire	Identification of social and economic background of respondents.
	(v) Opinionnaire	Use of attitude scales, projective techniques, use of sociometric scales.
	(vi) Personal interview	Interviewer uses a detailed schedule with open and closed questions.
	(vii) Focused interview	Interviewer focuses attention upon a given experience and its effects.
	(viii) Group interview	Small groups of respondents are interviewed simultaneously.
	(ix) Telephone survey	Used as a survey technique for information and for discerning opinion; may also be used as a follow up of questionnaire.
	(x) Case study and life history	Cross sectional collection of data for intensive analysis, longitudinal collection of data of intensive character.
3. Laboratory Research	Small group study of random behaviour, play and role analysis	Use of audio-visual recording devices, use of observers, etc.

From what has been stated above, we can say that methods are more general. It is the methods that generate techniques. However, in practice, the two terms are taken as interchangeable and when we talk of research methods we do, by implication, include research techniques within their compass.

use in performing research operations. In other words, all those methods which are used by the researcher during the course of studying his research problem are termed as research methods. Since the object of research, particularly the applied research, it to arrive at a solution for a given problem, the available data and the unknown aspects of the problem have to be related to each other to make a solution possible. Keeping this in view, research methods can be put into the following three groups:

1. In the first group we include those methods which are concerned with the collection of data. These methods will be used where the data already available are not sufficient to arrive at the required solution;
2. The second group consists of those statistical techniques which are used for establishing relationships between the data and the unknowns;
3. The third group consists of those methods which are used to evaluate the accuracy of the results obtained.

Research methods falling in the above stated last two groups are generally taken as the analytical tools of research.

Research methodology is a way to systematically solve the research problem. It may be understood as a science of studying how research is done scientifically. In it we study the various steps that are generally adopted by a researcher in studying his research problem along with the logic behind them. It is necessary for the researcher to know not only the research methods/techniques but also the methodology. Researchers not only need to know how to develop certain indices or tests, how to calculate the mean, the mode, the median or the standard deviation or chi-square, how to apply particular research techniques, but they also need to know which of these methods or techniques, are relevant and which are not, and what would they mean and indicate and why. Researchers also need to understand the assumptions underlying various techniques and they need to know the criteria by which they can decide that certain techniques and procedures will be applicable to certain problems and others will not. All this means that it is necessary for the researcher to design his methodology for his problem as the same may differ from problem to problem. For example, an architect, who designs a building, has to consciously evaluate the basis of his decisions, i.e., he has to evaluate why and on what basis he selects particular size, number and location of doors, windows and ventilators, uses particular materials and not others and the like. Similarly, in research the scientist has to expose the research decisions to evaluation before they are implemented. He has to specify very clearly and precisely what decisions he selects and why he selects them so that they can be evaluated by others also.

From what has been stated above, we can say that research methodology has many dimensions and research methods do constitute a part of the research methodology. The scope of research methodology is wider than that of research methods. *Thus, when we talk of research methodology we not only talk of the research methods but also consider the logic behind the methods we use in the context of our research study and explain why we are using a particular method or technique and why we are not using others so that research results are capable of being evaluated either by the researcher himself or by others*. Why a research study has been undertaken, how the research problem has been defined, in what way and why the hypothesis has been formulated, what data have been collected and what particular method has been adopted, why particular technique of analysing data has been used and a host of similar other questions are usually answered when we talk of research methodology concerning a research problem or study.

Research and Scientific Method

For a clear perception of the term research, one should know the meaning of scientific method. The two terms, research and scientific method, are closely related. Research, as we have already stated, can be termed as "an inquiry into the nature of, the reasons for, and the consequences of any particular set of circumstances, whether these circumstances are experimentally controlled or recorded just as they occur. Further, research implies the researcher is interested in more than particular results; he is interested in the repeatability of the results and in their extension to more complicated and general situations."[7] On the other hand, the philosophy common to all research methods and techniques, although they may vary considerably from one science to another, is usually given the name of scientific method. In this context, Karl Pearson writes, "The scientific method is one and same in the branches (of science) and that method is the method of all logically trained minds ... the unity of all sciences consists alone in its methods, not its material; the man who classifies facts of any kind whatever, who sees their mutual relation and describes their sequences, is applying the Scientific Method and is a man of science."[8] Scientific method is the pursuit of truth as determined by logical considerations. The ideal of science is to achieve a systematic interrelation of facts. Scientific method attempts to achieve "this ideal by experimentation, observation, logical arguments from accepted postulates and a combination of these three in varying proportions."[9] In scientific method, logic aids in formulating propositions explicitly and accurately so that their possible alternatives become clear. Further, logic develops the consequences of such alternatives, and when these are compared with observable phenomena, it becomes possible for the researcher or the scientist to state which alternative is most in harmony with the observed facts. All this is done through experimentation and survey investigations which constitute the integral parts of scientific method.

Experimentation is done to test hypotheses and to discover new relationships. If any, among variables. But the conclusions drawn on the basis of experimental data are generally criticized for either faulty assumptions, poorly designed experiments, badly executed experiments or faulty interpretations. As such the researcher must pay all possible attention while developing the experimental design and must state only probable inferences. The purpose of survey investigations may also be to provide scientifically gathered information to work as a basis for the researchers for their conclusions.

The scientific method is, thus, based on certain basic postulates which can be stated as under:

1. It relies on empirical evidence;
2. It utilizes relevant concepts;
3. It is committed to only objective considerations;
4. It presupposes ethical neutrality, i.e., it aims at nothing but making only adequate and correct statements about population objects;
5. It results into probabilistic predictions;
6. Its methodology is made known to all concerned for critical scrutiny are for use in testing the conclusions through replication;
7. It aims at formulating most general axioms or what can be termed as scientific theories.

[7] Bernard Ostle and Richard W. Mensing, *Statistics in Research*, p. 2

[8] Karl Pearson, *The Grammar of Science*, Part I, pp. 10–12.

[9] Ostle and Mensing: *op. cit.*, p. 2.

Thus, "the scientific method encourages a rigorous, impersonal mode of procedure dictated by the demands of logic and objective procedure."[10] Accordingly, scientific method implies an objective, logical and systematic method, i.e., a method free from personal bias or prejudice, a method to ascertain demonstrable qualities of a phenomenon capable of being verified, a method wherein the researcher is guided by the rules of logical reasoning, a method wherein the investigation proceeds in an orderly manner and a method that implies internal consistency.

Importance of Knowing How Research is Done

The study of research methodology gives the student the necessary training in gathering material and arranging or card-indexing them, participation in the field work when required, and also training in techniques for the collection of data appropriate to particular problems, in the use of statistics, questionnaires and controlled experimentation and in recording evidence, sorting it out and interpreting it. In fact, importance of knowing the methodology of research or how research is done stems from the following considerations:

(i) For one who is preparing himself for a career of carrying out research, the importance of knowing research methodology and research techniques is obvious since the same constitute the tools of his trade. The knowledge of methodology provides good training specially to the new research worker and enables him to do better research. It helps him to develop disciplined thinking or a 'bent of mind' to observe the field objectively. Hence, those aspiring for careerism in research must develop the skill of using research techniques and must thoroughly understand the logic behind them.

(ii) Knowledge of how to do research will inculcate the ability to evaluate and use research results with reasonable confidence. In other words, we can state that the knowledge of research methodology is helpful in various fields such as government or business administration, community development and social work where persons are increasingly called upon to evaluate and use research results for action.

(iii) When one knows how research is done, then one may have the satisfaction of acquiring a new intellectual tool which can become a way of looking at the world and of judging every day experience. Accordingly, it enables use to make intelligent decisions concerning problems facing us in practical life at different points of time. Thus, the knowledge of research methodology provides tools to took at things in life objectively.

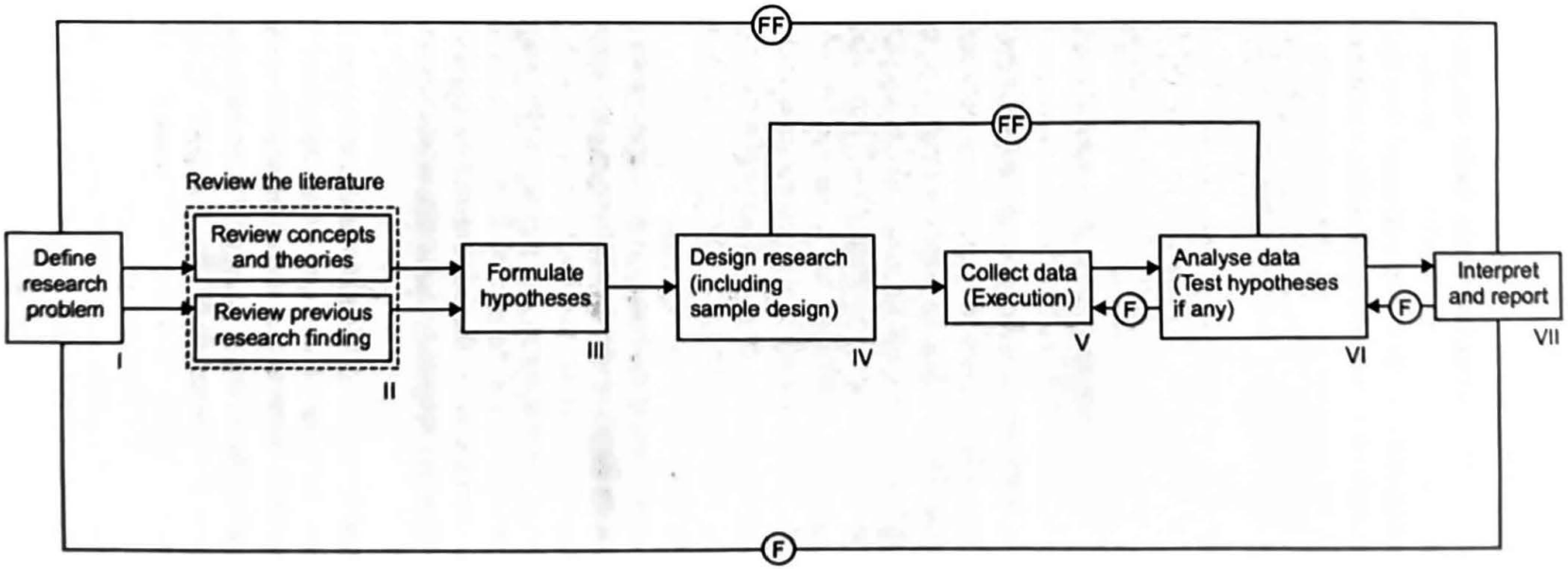
(iv) In this scientific age, all of us are in many ways consumers of research results and we can use them intelligently provided we are able to judge the adequacy of the methods by which they have been obtained. The knowledge of methodology helps the consumer of research results to evaluate them and enables him to take rational decisions.

Research Process

Before embarking on the details of research methodology and techniques, it seems appropriate to present a brief overview of the research process. Research process consists of series of actions or steps necessary to effectively carry out research and the desired sequencing of these steps. The chart shown in Figure 1.1 well illustrates a research process.

[10] Carlos L. Lastrucci, *The Scientific Approach: Basic Principles of the Scientific Method*, p. 7.

RESEARCH PROCESS IN FLOW CHART



Where (F) = feed back (Helps in controlling the sub-system to which it is transmitted)

(FF) = feed forward (Serves the vital function of providing criteria for evaluation)

Fig. 1.1

The chart indicates that the research process consists of a number of closely related activities, as shown through I to VII. But such activities overlap continuously rather than following a strictly prescribed sequence. At times, the first step determines the nature of the last step to be undertaken. If subsequent procedures have not been taken into account in the early stages, serious difficulties may arise which may even prevent the completion of the study. One should remember that the various steps involved in a research process are not mutually exclusive; nor they are separate and distinct. They do not necessarily follow each other in any specific order and the researcher has to be constantly anticipating at each step in the research process the requirements of the subsequent steps. However, the following order concerning various steps provides a useful procedural guideline regarding the research process: (1) formulating the research problem; (2) extensive literature survey; (3) developing the hypothesis; (4) preparing the research design; (5) determining sample design; (6) collecting the data; (7) execution of the project; (8) analysis of data; (9) hypothesis testing; (10) generalisations and interpretation, and (11) preparation of the report or presentation of the results, i.e., formal write-up of conclusions reached.

A brief description of the above stated steps will be helpful.

1. Formulating the research problem: There are two types of research problems, viz., those which relate to states of nature and those which relate to relationships between variables. At the very outset the researcher must single out the problem he wants to study, i.e., he must decide the general area of interest or aspect of a subject-matter that he would like to inquire into. Initially the problem may be stated in a broad general way and then the ambiguities, if any, relating to the problem be resolved. Then, the feasibility of a particular solution has to be considered before a working formulation of the problem can be set up. The formulation of a general topic into a specific research problem, thus, constitutes the first step in a scientific enquiry. Essentially two steps are involved in formulating the research problem, viz., understanding the problem thoroughly, and rephrasing the same into meaningful terms from an analytical point of view.

The best way of understanding the problem is to discuss it with one's own colleagues or with those having some expertise in the matter. In an academic institution the researcher can seek the help from a guide who is usually an experienced man and has several research problems in mind. Often, the guide puts forth the problem in general terms and it is up to the researcher to narrow it down and phrase the problem in operational terms. In private business units or in governmental organisations, the problem is usually earmarked by the administrative agencies with whom the researcher can discuss as to how the problem originally came about and what considerations are involved in its possible solutions.

The researcher must at the same time examine all available literature to get himself acquainted with the selected problem. He may review two types of literature—the conceptual literature concerning the concepts and theories, and the empirical literature consisting of studies made earlier which are similar to the one proposed. The basic outcome of this review will be the knowledge as to what data and other materials are available for operational purposes which will enable the researcher to specify his own research problem in a meaningful context. After this the researcher rephrases the problem into analytical or operational terms i.e., to put the problem in as specific terms as possible. This task of formulating, or defining, a research problem is a step of greatest importance in the entire research process. The problem to be investigated must be defined unambiguously for that will help discriminating relevant data from irrelevant ones. Care must, however, be taken to verify the objectivity and validity of the background facts concerning the problem. Professor W.A. Neiswanger correctly states that

the statement of the objective is of basic importance because it determines the data which are to be collected, the characteristics of the data which are relevant, relations which are to be explored, the choice of techniques to be used in these explorations and the form of the final report. If there are certain pertinent terms, the same should be clearly defined along with the task of formulating the problem. In fact, formulation of the problem often follows a sequential pattern where a number of formulations are set up, each formulation more specific than the preceeding one, each one phrased in more analytical terms, and each more realistic in terms of the available data and resources.

2. Extensive literature survey: Once the problem is formulated, a brief summary of it should be written down. It is compulsory for a research worker writing a thesis for a Ph.D. degree to write a synopsis of the topic and submit it to the necessary Committee or the Research Board for approval. At this juncture the researcher should undertake extensive literature survey connected with the problem. For this purpose, the abstracting and indexing journals and published or unpublished bibliographies are the first place to go to. Academic journals, conference proceedings, government reports, books etc., must be tapped depending on the nature of the problem. In this process, it should be remembered that one source will lead to another. The earlier studies, if any, which are similar to the study in hand should be carefully studied. A good library will be a great help to the researcher at this stage.

3. Development of working hypotheses: After extensive literature survey, researcher should state in clear terms the working hypothesis or hypotheses. Working hypothesis is tentative assumption made in order to draw out and test its logical or empirical consequences. As such the manner in which research hypotheses are developed is particularly important since they provide the focal point for research. They also affect the manner in which tests must be conducted in the analysis of data and indirectly the quality of data which is required for the analysis. In most types of research, the development of working hypothesis plays an important role. Hypothesis should be very specific and limited to the piece of research in hand because it has to be tested. The role of the hypothesis is to guide the researcher by delimiting the area of research and to keep him on the right track. It sharpens his thinking and focuses attention on the more important facets of the problem. It also indicates the type of data required and the type of methods of data analysis to be used.

How does one go about developing working hypotheses? The answer is by using the following approach:

(a) Discussions with colleagues and experts about the problem, its origin and the objectives in seeking a solution;
(b) Examination of data and records, if available, concerning the problem for possible trends, peculiarities and other clues;
(c) Review of similar studies in the area or of the studies on similar problems; and
(d) Exploratory personal investigation which involves original field interviews on a limited scale with interested parties and individuals with a view to secure greater insight into the practical aspects of the problem.

Thus, working hypotheses arise as a result of a-priori thinking about the subject, examination of the available data and material including related studies and the counsel of experts and interested parties. Working hypotheses are more useful when stated in precise and clearly defined terms. It may as well be remembered that occasionally we may encounter a problem where we do not need working

hypotheses, specially in the case of exploratory or formulative researches which do not aim at testing the hypothesis. But as a general rule, specification of working hypotheses in another basic step of the research process in most research problems.

4. Preparing the research design: The research problem having been formulated in clear cut terms, the researcher will be required to prepare a research design, i.e., he will have to state the conceptual structure within which research would be conducted. The preparation of such a design facilitates research to be as efficient as possible yielding maximal information. In other words, the function of research design is to provide for the collection of relevant evidence with minimal expenditure of effort, time and money. But how all these can be achieved depends mainly on the research purpose. Research purposes may be grouped into four categories, viz., (i) Exploration, (ii) Description, (iii) Diagnosis, and (iv) Experimentation. A flexible research design which provides opportunity for considering many different aspects of a problem is considered appropriate if the purpose of the research study is that of exploration. But when the purpose happens to be an accurate description of a situation or of an association between variables, the suitable design will be one that minimises bias and maximises the reliability of the data collected and analysed.

There are several research designs, such as, experimental and non-experimental hypothesis testing. Experimental designs can be either informal designs (such as before-and-after without control, after-only with control, before-and-after with control) or formal designs (such as completely randomized design, randomized block design, Latin square design, simple and complex factorial designs), out of which the researcher must select one for his own project.

The preparation of the research design, appropriate for a particular research problem, involves usually the consideration of the following:

(i) the means of obtaining the information;
(ii) the availability and skills of the researcher and his staff (if any);
(iii) explanation of the way in which selected means of obtaining information will be organised and the reasoning leading to the selection;
(iv) the time available for research; and
(v) the cost factor relating to research, i.e., the finance available for the purpose.

5. Determining sample design: All the items under consideration in any field of inquiry constitute a 'universe' or 'population'. A complete enumeration of all the items in the 'population' is known as a census inquiry. It can be presumed that in such an inquiry when all the items are covered no element of chance is left and highest accuracy is obtained. But in practice this may not be true. Even the slightest element of bias in such an inquiry will get larger and larger as the number of observations increases. Moreover, there is no way of checking the element of bias or its extent except through a resurvey or use of sample checks. Besides, this type of inquiry involves a great deal of time, money and energy. Not only this, census inquiry is not possible in practice under many circumstances. For instance, blood testing is done only on sample basis. Hence, quite often we select only a few items from the universe for our study purposes. The items so selected constitute what is technically called a sample.

The researcher must decide the way of selecting a sample or what is popularly known as the sample design. In other words, a sample design is a definite plan determined before any data are actually collected for obtaining a sample from a given population. Thus, the plan to select 12 of a

city's 200 drugstores in a certain way constitutes a sample design. Samples can be either probability samples or non-probability samples. With probability samples each element has a known probability of being included in the sample but the non-probability samples do not allow the researcher to determine this probability. Probability samples are those based on simple random sampling, systematic sampling, stratified sampling, cluster/area sampling whereas non-probability samples are those based on convenience sampling, judgement sampling and quota sampling techniques. A brief mention of the important sample designs is as follows:

(i) *Deliberate sampling:* Deliberate sampling is also known as purposive or non-probability sampling. This sampling method involves purposive or deliberate selection of particular units of the universe for constituting a sample which represents the universe. When population elements are selected for inclusion in the sample based on the ease of access, it can be called *convenience sampling.* If a researcher wishes to secure data from, say, gasoline buyers, he may select a fixed number of petrol stations and may conduct interviews at these stations. This would be an example of convenience sample of gasoline buyers. At times such a procedure may give very biased results particularly when the population is not homogeneous. On the other hand, in *judgement sampling* the researcher's judgement is used for selecting items which he considers as representative of the population. For example, a judgement sample of college students might be taken to secure reactions to a new method of teaching. Judgement sampling is used quite frequently in qualitative research where the desire happens to be to develop hypotheses rather than to generalise to larger populations.

(ii) *Simple random sampling:* This type of sampling is also known as chance sampling or probability sampling where each and every item in the population has an equal chance of inclusion in the sample and each one of the possible samples, in case of finite universe, has the same probability of being selected. For example, if we have to select a sample of 300 items from a universe of 15,000 items, then we can put the names or numbers of all the 15,000 items on slips of paper and conduct a lottery. Using the random number tables is another method of random sampling. To select the sample, each item is assigned a number from 1 to 15,000. Then, 300 five digit random numbers are selected from the table. To do this we select some random starting point and then a systematic pattern is used in proceeding through the table. We might start in the 4th row, second column and proceed down the column to the bottom of the table and then move to the top of the next column to the right. When a number exceeds the limit of the numbers in the frame, in our case over 15,000, it is simply passed over and the next number selected that does fall within the relevant range. Since the numbers were placed in the table in a completely random fashion, the resulting sample is random. This procedure gives each item an equal probability of being selected. In case of infinite population, the selection of each item in a random sample is controlled by the same probability and that successive selections are independent of one another.

(iii) *Systematic sampling:* In some instances the most practical way of sampling is to select every 15th name on a list, every 10th house on one side of a street and so on. Sampling of this type is known as systematic sampling. An element of randomness is usually introduced into this kind of sampling by using random numbers to pick up the unit with which to start. This procedure is useful when sampling frame is available in the form of a list. In such a design the selection process starts by picking some random point in the list and then every *n*th element is selected until the desired number is secured.

(iv) *Stratified sampling:* If the population from which a sample is to be drawn does not constitute a homogeneous group, then stratified sampling technique is applied so as to obtain a representative sample. In this technique, the population is stratified into a number of non-overlapping subpopulations or strata and sample items are selected from each stratum. If the items selected from each stratum is based on simple random sampling the entire procedure, first stratification and then simple random sampling, is known as *stratified random sampling.*

(v) *Quota sampling:* In stratified sampling the cost of taking random samples from individual strata is often so expensive that interviewers are simply given quota to be filled from different strata, the actual selection of items for sample being left to the interviewer's judgement. This is called quota sampling. The size of the quota for each stratum is generally proportionate to the size of that stratum in the population. Quota sampling is thus an important form of non-probability sampling. Quota samples generally happen to be judgement samples rather than random samples.

(vi) *Cluster sampling and area sampling:* Cluster sampling involves grouping the population and then selecting the groups or the clusters rather than individual elements for inclusion in the sample. Suppose some departmental store wishes to sample its credit card holders. It has issued its cards to 15,000 customers. The sample size is to be kept say 450. For cluster sampling this list of 15,000 card holders could be formed into 100 clusters of 150 card holders each. Three clusters might then be selected for the sample randomly. The sample size must often be larger than the simple random sample to ensure the same level of accuracy because is cluster sampling procedural potential for order bias and other sources of error is usually accentuated. The clustering approach can, however, make the sampling procedure relatively easier and increase the efficiency of field work, specially in the case of personal interviews.

Area sampling is quite close to cluster sampling and is often talked about when the total geographical area of interest happens to be big one. Under area sampling we first divide the total area into a number of smaller non-overlapping areas, generally called geographical clusters, then a number of these smaller areas are randomly selected, and all units in these small areas are included in the sample. Area sampling is specially helpful where we do not have the list of the population concerned. It also makes the field interviewing more efficient since interviewer can do many interviews at each location.

(vii) *Multi-stage sampling:* This is a further development of the idea of cluster sampling. This technique is meant for big inquiries extending to a considerably large geographical area like an entire country. Under multi-stage sampling the first stage may be to select large primary sampling units such as states, then districts, then towns and finally certain families within towns. If the technique of random-sampling is applied at all stages, the sampling procedure is described as multi-stage random sampling.

(viii) *Sequential sampling:* This is somewhat a complex sample design where the ultimate size of the sample is not fixed in advance but is determined according to mathematical decisions on the basis of information yielded as survey progresses. This design is usually adopted under acceptance sampling plan in the context of statistical quality control.

In practice, several of the methods of sampling described above may well be used in the same study in which case it can be called mixed sampling. It may be pointed out here that normally one

should resort to random sampling so that bias can be eliminated and sampling error can be estimated. But purposive sampling is considered desirable when the universe happens to be small and a known characteristic of it is to be studied intensively. Also, there are conditions under which sample designs other than random sampling may be considered better for reasons like convenience and low costs. *The sample design to be used must be decided by the researcher taking into consideration the nature of the inquiry and other related factors.*

6. Collecting the data: In dealing with any real life problem it is often found that data at hand are inadequate, and hence, it becomes necessary to collect data that are appropriate. There are several ways of collecting the appropriate data which differ considerably in context of money costs, time and other resources at the disposal of the researcher.

Primary data can be collected either through experiment or through survey. If the researcher conducts an experiment, he observes some quantitative measurements, or the data, with the help of which he examines the truth contained in his hypothesis. But in the case of a survey, data can be collected by any one or more of the following ways:

(i) *By observation:* This method implies the collection of information by way of investigator's own observation, without interviewing the respondents. The information obtained relates to what is currently happening and is not complicated by either the past behaviour or future intentions or attitudes of respondents. This method is no doubt an expensive method and the information provided by this method is also very limited. As such this method is not suitable in inquiries where large samples are concerned.

(ii) *Through personal interview:* The investigator follows a rigid procedure and seeks answers to a set of pre-conceived questions through personal interviews. This method of collecting data is usually carried out in a structured way where output depends upon the ability of the interviewer to a large extent.

(iii) *Through telephone interviews:* This method of collecting information involves contacting the respondents on telephone itself. This is not a very widely used method but it plays an important role in industrial surveys in developed regions, particularly, when the survey has to be accomplished in a very limited time.

(iv) *By mailing of questionnaires:* The researcher and the respondents do come in contact with each other if this method of survey is adopted. Questionnaires are mailed to the respondents with a request to return after completing the same. It is the most extensively used method in various economic and business surveys. Before applying this method, usually a Pilot Study for testing the questionnaire is conduced which reveals the weaknesses, if any, of the questionnaire. Questionnaire to be used must be prepared very carefully so that it may prove to be effective in collecting the relevant information.

(v) *Through schedules:* Under this method the enumerators are appointed and given training. They are provided with schedules containing relevant questions. These enumerators go to respondents with these schedules. Data are collected by filling up the schedules by enumerators on the basis of replies given by respondents. Much depends upon the capability of enumerators so far as this method is concerned. Some occasional field checks on the work of the enumerators may ensure sincere work.

The researcher should select one of these methods of collecting the data taking into consideration the nature of investigation, objective and scope of the inquiry, finanical resources, available time and the desired degree of accuracy. Though he should pay attention to all these factors but much depends upon the ability and experience of the researcher. In this context *Dr A.L. Bowley* very aptly remarks that in collection of statistical data commonsense is the chief requisite and experience the chief teacher.

7. Execution of the project: Execution of the project is a very important step in the research process. If the execution of the project proceeds on correct lines, the data to be collected would be adequate and dependable. The researcher should see that the project is executed in a systematic manner and in time. If the survey is to be conducted by means of structured questionnaires, data can be readily machine-processed. In such a situation, questions as well as the possible answers may be coded. If the data are to be collected through interviewers, arrangements should be made for proper selection and training of the interviewers. The training may be given with the help of instruction manuals which explain clearly the job of the interviewers at each step. Occasional field checks should be made to ensure that the interviewers are doing their assigned job sincerely and efficiently. A careful watch should be kept for unanticipated factors in order to keep the survey as much realistic as possible. This, in other words, means that steps should be taken to ensure that the survey is under statistical control so that the collected information is in accordance with the pre-defined standard of accuracy. If some of the respondents do not cooperate, some suitable methods should be designed to tackle this problem. One method of dealing with the non-response problem is to make a list of the non-respondents and take a small sub-sample of them, and then with the help of experts vigorous efforts can be made for securing response.

8. Analysis of data: After the data have been collected, the researcher turns to the task of analysing them. The analysis of data requires a number of closely related operations such as establishment of categories, the application of these categories to raw data through coding, tabulation and then drawing statistical inferences. The unwieldy data should necessarily be condensed into a few manageable groups and tables for further analysis. Thus, researcher should classify the raw data into some purposeful and usable categories. *Coding* operation is usually done at this stage through which the categories of data are transformed into symbols that may be tabulated and counted. *Editing* is the procedure that improves the quality of the data for coding. With coding the stage is ready for tabulation. *Tabulation* is a part of the technical procedure wherein the classified data are put in the form of tables. The mechanical devices can be made use of at this juncture. A great deal of data, specially in large inquiries, is tabulated by computers. Computers not only save time but also make it possible to study large number of variables affecting a problem simultaneously.

Analysis work after tabulation is generally based on the computation of various percentages, coefficients, etc., by applying various well defined statistical formulae. In the process of analysis, relationships or differences supporting or conflicting with original or new hypotheses should be subjected to tests of significance to determine with what validity data can be said to indicate any conclusion(s). For instance, if there are two samples of weekly wages, each sample being drawn from factories in different parts of the same city, giving two different mean values, then our problem may be whether the two mean values are significantly different or the difference is just a matter of chance. Through the use of statistical tests we can establish whether such a difference is a real one or is the result of random fluctuations. If the difference happens to be real, the inference will be that the two samples

come from different universes and if the difference is due to chance, the conclusion would be that the two samples belong to the same universe. Similarly, the technique of analysis of variance can help us in analysing whether three or more varieties of seeds grown on certain fields yield significantly different results or not. In brief, the researcher can analyse the collected data with the help of various statistical measures.

9. Hypothesis-testing: After analysing the data as stated above, the researcher is in a position to test the hypotheses, if any, he had formulated earlier. Do the facts support the hypotheses or they happen to be contrary? This is the usual question which should be answered while testing hypotheses. Various tests, such as Chi square test, *t*-test, *F*-test, have been developed by statisticians for the purpose. The hypotheses may be tested through the use of one or more of such tests, depending upon the nature and object of research inquiry. Hypothesis-testing will result in either accepting the hypothesis or in rejecting it. If the researcher had no hypotheses to start with, generalisations established on the basis of data may be stated as hypotheses to be tested by subsequent researches in times to come.

10. Generalisations and interpretation: If a hypothesis is tested and upheld several times, it may be possible for the researcher to arrive at generalisation, i.e., to build a theory. As a matter of fact, the real value of research lies in its ability to arrive at certain generalisations. If the researcher had no hypothesis to start with, he might seek to explain his findings on the basis of some theory. It is known as interpretation. The process of interpretation may quite often trigger off new questions which in turn may lead to further researches.

11. Preparation of the report or the thesis: Finally, the researcher has to prepare the report of what has been done by him. Writing of report must be done with great care keeping in view the following:

1. The layout of the report should be as follows: (*i*) the preliminary pages; (*ii*) the main text, and (*iii*) the end matter.

   In its preliminary pages the report should carry title and date followed by acknowledgements and foreword. Then there should be a table of contents followed by a list of tables and list of graphs and charts, if any, given in the report.

   The main text of the report should have the following parts:

   (a) *Introduction:* It should contain a clear statement of the objective of the research and an explanation of the methodology adopted in accomplishing the research. The scope of the study along with various limitations should as well be stated in this part.

   (b) *Summary of findings:* After introduction there would appear a statement of findings and recommendations in non-technical language. If the findings are extensive, they should be summarised.

   (c) *Main report:* The main body of the report should be presented in logical sequence and broken-down into readily identifiable sections.

   (d) *Conclusion:* Towards the end of the main text, researcher should again put down the results of his research clearly and precisely. In fact, it is the final summing up.

At the end of the report, appendices should be enlisted in respect of all technical data. Bibliography, i.e., list of books, journals, reports, etc., consulted, should also be given in the end. Index should also be given specially in a published research report.

2. Report should be written in a concise and objective style in simple language avoiding vague expressions such as 'it seems,' 'there may be', and the like.
3. Charts and illustrations in the main report should be used only if they present the information more clearly and forcibly.
4. Calculated 'confidence limits' must be mentioned and the various constraints experienced in conducting research operations may as well be stated.

Criteria of Good Research

Whatever may be the types of research works and studies, one thing that is important is that they all meet on the common ground of scientific method employed by them. One expects scientific research to satisfy the following criteria:[11]

1. The purpose of the research should be clearly defined and common concepts be used.
2. The research procedure used should be described in sufficient detail to permit another researcher to repeat the research for further advancement, keeping the continuity of what has already been attained.
3. The procedural design of the research should be carefully planned to yield results that are as objective as possible.
4. The researcher should report with complete frankness, flaws in procedural design and estimate their effects upon the findings.
5. The analysis of data should be sufficiently adequate to reveal its significance and the methods of analysis used should be appropriate. The validity and reliability of the data should be checked carefully.
6. Conclusions should be confined to those justified by the data of the research and limited to those for which the data provide an adequate basis.
7. Greater confidence in research is warranted if the researcher is experienced, has a good reputation in research and is a person of integrity.

In other words, we can state the qualities of a good research[12] as under:

1. *Good research is systematic:* It means that research is structured with specified steps to be taken in a specified sequence in accordance with the well defined set of rules. Systematic characteristic of the research does not rule out creative thinking but it certainly does reject the use of guessing and intuition in arriving at conclusions.
2. *Good research is logical:* This implies that research is guided by the rules of logical reasoning and the logical process of induction and deduction are of great value in carrying out research. Induction is the process of reasoning from a part to the whole whereas deduction is the process of reasoning from some premise to a conclusion which follows from that very premise. In fact, logical reasoning makes research more meaningful in the context of decision making.

[11] James Harold Fox, Criteria of Good Research, Phi Delta Kappan, Vol. 39 (March, 1958), pp. 285–86.

[12] See, Danny N. Bellenger and Barnett, A. Greenberg, "*Marketing Research—A Management Information Approach*", p. 107–108.

3. *Good research is empirical:* It implies that research is related basically to one or more aspects of a real situation and deals with concrete data that provides a basis for external validity to research results.
4. *Good research is replicable:* This characteristic allows research results to be verified by replicating the study and thereby building a sound basis for decisions.

Problems Encountered by Researchers in India

Researchers in India, particularly those engaged in empirical research, are facing several problems. Some of the important problems are as follows:

1. *The lack of a scientific training in the methodology of research* is a great impediment for researchers in our country. There is paucity of competent researchers. Many researchers take a leap in the dark without knowing research methods. Most of the work, which goes in the name of research is not methodologically sound. Research to many researchers and even to their guides, is mostly a scissor and paste job without any insight shed on the collated materials. The consequence is obvious, viz., the research results, quite often, do not reflect the reality or realities. Thus, a systematic study of research methodology is an urgent necessity. Before undertaking research projects, researchers should be well equipped with all the methodological aspects. As such, *efforts should be made to provide short-duration intensive courses for meeting this requirement.*
2. There is *insufficient interaction* between the university research departments on one side and business establishments, government departments and research institutions on the other side. A great deal of primary data of non-confidential nature remain untouched/untreated by the researchers for want of proper contacts. *Efforts should be made to develop satisfactory liaison among all concerned for better and realistic researches.* There is need for developing some mechanisms of a university—industry interaction programme so that academics can get ideas from practitioners on what needs to be researched and practitioners can apply the research done by the academics.
3. Most of the business units in our country do not have the confidence that the material supplied by them to researchers will not be misused and as such they are often reluctant in supplying the needed information to researchers. The concept of secrecy seems to be sacrosanct to business organisations in the country so much so that it proves an impermeable barrier to researchers. Thus, *there is the need for generating the confidence that the information/data obtained from a business unit will not be misused.*
4. *Research studies overlapping one another are undertaken quite often for want of adequate information.* This results in duplication and fritters away resources. This problem can be solved by proper compilation and revision, at regular intervals, of a list of subjects on which and the places where the research is going on. Due attention should be given toward identification of research problems in various disciplines of applied science which are of immediate concern to the industries.
5. *There does not exist a code of conduct for researchers* and inter-university and inter-departmental rivalries are also quite common. Hence, there is need for developing a code of conduct for researchers which, if adhered sincerely, can win over this problem.

6. Many researchers in our country also face *the difficulty of adequate and timely secretarial assistance,* including computerial assistance. This causes unnecessary delays in the completion of research studies. All possible efforts be made in this direction so that efficient secretarial assistance is made available to researchers and that too well in time. University Grants Commission must play a dynamic role in solving this difficulty.
7. *Library management and functioning is not satisfactory at many places* and much of the time and energy of researchers are spent in tracing out the books, journals, reports, etc., rather than in tracing out relevant material from them.
8. *There is also the problem that many of our libraries are not able to get copies of old and new Acts/Rules, reports and other government publications in time.* This problem is felt more in libraries which are away in places from Delhi and/or the state capitals. Thus, efforts should be made for the regular and speedy supply of all governmental publications to reach our libraries.
9. *There is also the difficulty of timely availability of published data* from various government and other agencies doing this job in our country. Researcher also faces the problem on account of the fact that the published data vary quite significantly because of differences in coverage by the concerning agencies.
10. There may, at times, take place *the problem of conceptualization* and also problems relating to the process of data collection and related things.

Questions

1. Briefly describe the different steps involved in a research process.
2. What do you mean by research? Explain its significance in modern times.
3. Distinguish between Research methods and Research methodology.
4. Describe the different types of research, clearly pointing out the difference between an experiment and a survey.
5. Write short notes on:
   (1) Design of the research project;
   (2) Ex post facto research;
   (3) Motivation in research;
   (4) Objectives of research;
   (5) Criteria of good research;
   (7) Research and scientific method.
6. "Empirical research in India in particular creates so many problems for the researchers". State the problems that are usually faced by such researchers.

   (Raj. Univ. EAFM., M. Phil. Exam., 1979)
7. "A research scholar has to work as a judge and derive the truth and not as a pleader who is only eager to prove his case in favour of his plaintiff." Discuss the statement pointing out the objectives of research.

8. "Creative management, whether in public administration or private industry, depends on methods of inquiry that maintain objectivity, clarity, accuracy and consistency". Discuss this statement and examine the significance of research".

(Raj. Univ. EAFM., M. Phil. Exam., 1978)

9. "Research is much concerned with proper fact finding, analysis and evaluation." Do you agree with this statement? Give reasons in support of your answer.

10. It is often said that there is not a proper link between some of the activities under way in the world of academics and in most business in our country. Account for this state of affairs and give suggestions for improvement.

2

Defining the Research Problem

In research process, the first and foremost step happens to be that of selecting and properly defining a research problem.* A researcher must find the problem and formulate it so that it becomes susceptible to research. Like a medical doctor, a researcher must examine all the symptoms (presented to him or observed by him) concerning a problem before he can diagnose correctly. To define a problem correctly, a researcher must know: what a problem is?

WHAT IS A RESEARCH PROBLEM?

A research problem, in general, refers to some difficulty which a researcher experiences in the context of either a theoretical or practical situation and wants to obtain a solution for the same. Usually we say that a research problem does exist if the following conditions are met with:

(i) There must be an individual (or a group or an organisation), let us call it 'I,' to whom the problem can be attributed. The individual or the organisation, as the case may be, occupies an environment, say 'N', which is defined by values of the uncontrolled variables, Y_j.

(ii) There must be at least two courses of action, say C_1 and C_2, to be pursued. A course of action is defined by one or more values of the controlled variables. For example, the number of items purchased at a specified time is said to be one course of action.

(iii) There must be at least two possible outcomes, say O_1 and O_2, of the course of action, of which one should be preferable to the other. In other words, this means that there must be at least one outcome that the researcher wants, i.e., an objective.

(iv) The courses of action available must provides some chance of obtaining the objective, but they cannot provide the same chance, otherwise the choice would not matter. Thus, if $P(O_j \mid I, C_j, N)$ represents the probability that an outcome O_j will occur, if I select C_j in N, then $P(O_1 \mid I, C_1, N) \neq P(O_1 \mid I, C_2, N)$. In simple words, we can say that the choices must have unequal efficiencies for the desired outcomes.

* We talk of a research problem or hypothesis in case of descriptive or hypothesis testing research studies. Exploratory or formulative research studies do not start with a problem or hypothesis, their problem is to find a problem or the hypothesis to be tested. One should make a clear statement to this effect. This aspect has been dealt with in chapter entitled "Research Design"

Over and above these conditions, the individual or the organisation can be said to have the problem only if '*I*' does not know what course of action is best, i.e., '*I*', must be in doubt about the solution. Thus, an individual or a group of persons can be said to have a problem which can be technically described as a research problem, if they (individual or the group), having one or more desired outcomes, are confronted with two or more courses of action that have some but not equal efficiency for the desired objective(s) and are in doubt about which course of action is best.

We can, thus, state the components[1] of a research problem as under:

(i) There must be an individual or a group which has some difficulty or the problem.

(ii) There must be some objective(s) to be attained at. If one wants nothing, one cannot have a problem.

(iii) There must be alternative means (or the courses of action) for obtaining the objective(s) one wishes to attain. This means that there must be *at least two means* available to a researcher for if he has no choice of means, he cannot have a problem.

(iv) There must remain some doubt in the mind of a researcher with regard to the selection of alternatives. This means that research must answer the question concerning the relative efficiency of the possible alternatives.

(v) There must be some environment(s) to which the difficulty pertains.

Thus, a research problem is one which requires a researcher to find out the best solution for the given problem, i.e., to find out by which course of action the objective can be attained optimally in the context of a given environment. There are several factors which may result in making the problem complicated. For instance, the environment may change affecting the efficiencies of the courses of action or the values of the outcomes; the number of alternative courses of action may be very large; persons not involved in making the decision may be affected by it and react to it favourably or unfavourably, and similar other factors. All such elements (or at least the important ones) may be thought of in context of a research problem.

SELECTING THE PROBLEM

The research problem undertaken for study must be carefully selected. The task is a difficult one, although it may not appear to be so. Help may be taken from a research guide in this connection. Nevertheless, every researcher must find out his own salvation for research problems cannot be borrowed. A problem must spring from the researcher's mind like a plant springing from its own seed. If our eyes need glasses, it is not the optician alone who decides about the number of the lens we require. We have to see ourselves and enable him to prescribe for us the right number by cooperating with him. Thus, a research guide can at the most only help a researcher choose a subject. However, the following points may be observed by a researcher in selecting a research problem or a subject for research:

(i) Subject which is overdone should not be normally chosen, for it will be a difficult task to throw any new light in such a case.

(ii) Controversial subject should not become the choice of an average researcher.

[1] R.L. Ackoff, *The Design of Social Research*, Chicago University Press, Chicago, 1961.

(iii) Too narrow or too vague problems should be avoided.

(iv) The subject selected for research should be familiar and feasible so that the related research material or sources of research are within one's reach. Even then it is quite difficult to supply definitive ideas concerning how a researcher should obtain ideas for his research. For this purpose, a researcher should contact an expert or a professor in the University who is already engaged in research. He may as well read articles published in current literature available on the subject and may think how the techniques and ideas discussed therein might be applied to the solution of other problems. He may discuss with others what he has in mind concerning a problem. In this way he should make all possible efforts in selecting a problem.

(v) The importance of the subject, the qualifications and the training of a researcher, the costs involved, the time factor are few other criteria that must also be considered in selecting a problem. In other words, before the final selection of a problem is done, a researcher must ask himself the following questions:

(a) Whether he is well equipped in terms of his background to carry out the research?

(b) Whether the study falls within the budget he can afford?

(c) Whether the necessary cooperation can be obtained from those who must participate in research as subjects?

If the answers to all these questions are in the affirmative, one may become sure so far as the practicability of the study is concerned.

(vi) The selection of a problem must be preceded by a preliminary study. This may not be necessary when the problem requires the conduct of a research closely similar to one that has already been done. But when the field of inquiry is relatively new and does not have available a set of well developed techniques, a brief feasibility study must always be undertaken.

If the subject for research is selected properly by observing the above mentioned points, the research will not be a boring drudgery, rather it will be love's labour. In fact, zest for work is a must. The subject or the problem selected must involve the researcher and must have an upper most place in his mind so that he may undertake all pains needed for the study.

NECESSITY OF DEFINING THE PROBLEM

Quite often we all hear that a problem clearly stated is a problem half solved. This statement signifies the need for defining a research problem. The problem to be investigated must be defined unambiguously for that will help to discriminate relevant data from the irrelevant ones. A proper definition of research problem will enable the researcher to be on the track whereas an ill-defined problem may create hurdles. Questions like: What data are to be collected? What characteristics of data are relevant and need to be studied? What relations are to be explored. What techniques are to be used for the purpose? and similar other questions crop up in the mind of the researcher who can well plan his strategy and find answers to all such questions only when the research problem has been well defined. Thus, defining a research problem properly is a prerequisite for any study and is a step of the highest importance. In fact, formulation of a problem is often more essential than its

solution. It is only on careful detailing the research problem that we can work out the research design and can smoothly carry on all the consequential steps involved while doing research.

TECHNIQUE INVOLVED IN DEFINING A PROBLEM

Let us start with the question: What does one mean when he/she wants to define a research problem? The answer may be that one wants to state the problem along with the bounds within which it is to be studied. In other words, defining a problem involves the task of laying down boundaries within which a researcher shall study the problem with a pre-determined objective in view.

How to define a research problem is undoubtedly a herculean task. However, it is a task that must be tackled intelligently to avoid the perplexity encountered in a research operation. The usual approach is that the researcher should himself pose a question (or in case someone else wants the researcher to carry on research, the concerned individual, organisation or an authority should pose the question to the researcher) and set-up techniques and procedures for throwing light on the question concerned for formulating or defining the research problem. But such an approach generally does not produce definitive results because the question phrased in such a fashion is usually in broad general terms and as such may not be in a form suitable for testing.

Defining a research problem properly and clearly is a crucial part of a research study and must in no case be accomplished hurriedly. However, in practice this a frequently overlooked which causes a lot of problems later on. Hence, the research problem should be defined in a systematic manner, giving due weightage to all relating points. The technique for the purpose involves the undertaking of the following steps generally one after the other: (i) statement of the problem in a general way; (ii) understanding the nature of the problem; (iii) surveying the available literature (iv) developing the ideas through discussions; and (v) rephrasing the research problem into a working proposition.

A brief description of all these points will be helpful.

(i) Statement of the problem in a general way: First of all the problem should be stated in a broad general way, keeping in view either some practical concern or some scientific or intellectual interest. For this purpose, the researcher must immerse himself thoroughly in the subject matter concerning which he wishes to pose a problem. In case of social research, it is considered advisable to do some field observation and as such the researcher may undertake some sort of preliminary survey or what is often called *pilot survey*. Then the researcher can himself state the problem or he can seek the guidance of the guide or the subject expert in accomplishing this task. Often, the guide puts forth the problem in general terms, and it is then up to the researcher to narrow it down and phrase the problem in operational terms. In case there is some directive from an organisational authority, the problem then can be stated accordingly. The problem stated in a broad general way may contain various ambiguities which must be resolved by cool thinking and rethinking over the problem. At the same time the feasibility of a particular solution has to be considered and the same should be kept in view while stating the problem.

(ii) Understanding the nature of the problem: The next step in defining the problem is to understand its origin and nature clearly. The best way of understanding the problem is to discuss it with those who first raised it in order to find out how the problem originally came about and with what objectives in view. If the researcher has stated the problem himself, he should consider once again all those points that induced him to make a general statement concerning the problem. For a better

understanding of the nature of the problem involved, he can enter into discussion with those who have a good knowledge of the problem concerned or similar other problems. The researcher should also keep in view the environment within which the problem is to be studied and understood.

(iii) Surveying the available literature: All available literature concerning the problem at hand must necessarily be surveyed and examined before a definition of the research problem is given. This means that the researcher must be well-conversant with relevant theories in the field, reports and records as also all other relevant literature. He must devote sufficient time in reviewing of research already undertaken on related problems. This is done to find out what data and other materials, if any, are available for operational purposes. "Knowing what data are available often serves to narrow the problem itself as well as the technique that might be used."[2]. This would also help a researcher to know if there are certain gaps in the theories, or whether the existing theories applicable to the problem under study are inconsistent with each other, or whether the findings of the different studies do not follow a pattern consistent with the theoretical expectations and so on. All this will enable a researcher to take new strides in the field for furtherance of knowledge i.e., he can move up starting from the existing premise. Studies on related problems are useful for indicating the type of difficulties that may be encountered in the present study as also the possible analytical shortcomings. At times such studies may also suggest useful and even new lines of approach to the present problem.

(iv) Developing the ideas through discussions: Discussion concerning a problem often produces useful information. Various new ideas can be developed through such an exercise. Hence, a researcher must discuss his problem with his colleagues and others who have enough experience in the same area or in working on similar problems. This is quite often known as an *experience survey*. People with rich experience are in a position to enlighten the researcher on different aspects of his proposed study and their advice and comments are usually invaluable to the researcher. They help him sharpen his focus of attention on specific aspects within the field. Discussions with such persons should not only be confined to the formulation of the specific problem at hand, but should also be concerned with the general approach to the given problem, techniques that might be used, possible solutions, etc.

(v) Rephrasing the research problem: Finally, the researcher must sit to rephrase the research problem into a working proposition. Once the nature of the problem has been clearly understood, the environment (within which the problem has got to be studied) has been defined, discussions over the problem have taken place and the available literature has been surveyed and examined, rephrasing the problem into analytical or operational terms is not a difficult task. Through rephrasing, the researcher puts the research problem in as specific terms as possible so that it may become operationally viable and may help in the development of working hypotheses.*

In addition to what has been stated above, the following points must also be observed while defining a research problem:

[2] Robert Ferber and P.J. Verdoorn, *Research Methods in Economics and Business*, p. 33–34.

* Working hypotheses are a set of suggested tentative solutions of explanations of a research problem which may or may not be the real solutions. The task of research is to test and establish such hypotheses. Hypotheses should be clearly and precisely stated in simple terms, they should be testable, limited in scope and should state relationship between variables. They should be amenable to testing within a reasonable time and should be consistent with most of the known facts (Testing of hypotheses has been dealt with later in the book).

(a) Technical terms and words or phrases, with special meanings used in the statement of the problem, should be clearly defined.

(b) Basic assumptions or postulates (if any) relating to the research problem should be clearly stated.

(c) A straight forward statement of the value of the investigation (i.e., the criteria for the selection of the problem) should be provided.

(d) The suitability of the time-period and the sources of data available must also be considered by the researcher in defining the problem.

(e) The scope of the investigation or the limits within which the problem is to be studied must be mentioned explicitly in defining a research problem.

AN ILLUSTRATION

The technique of defining a problem outlined above can be illustrated for better understanding by taking an example as under:

Let us suppose that a research problem in a broad general way is as follows:

> "Why is productivity in Japan so much higher than in India"?
>
> In this form the question has a number of ambiguities such as: What sort of productivity is being referred to? With what industries the same is related? With what period of time the productivity is being talked about? In view of all such ambiguities the given statement or the question is much too general to be amenable to analysis. Rethinking and discussions about the problem may result in narrowing down the question to:
>
> "What factors were responsible for the higher labour productivity of Japan's manufacturing industries during the decade 1971 to 1980 relative to India's manufacturing industries?"
>
> This latter version of the problem is definitely an improvement over its earlier version for the various ambiguities have been removed to the extent possible. Further rethinking and rephrasing might place the problem on a still better operational basis as shown below:
>
> "To what extent did labour productivity in 1971 to 1980 in Japan exceed that of India in respect of 15 selected manufacturing industries? What factors were responsible for the productivity differentials between the two countries by industries?"

With this sort of formulation, the various terms involved such as 'labour productivity', 'productivity differentials', etc. must be explained clearly. The researcher must also see that the necessary data are available. In case the data for one or more industries selected are not available for the concerning time-period, then the said industry or industries will have to be substituted by other industry or industries. The suitability of the time-period must also be examined. Thus, all relevant factors must be considered by a researcher before finally defining a research problem.

CONCLUSION

We may conclude by saying that the task of defining a research problem, very often, follows a sequential pattern—the problem is stated in a general way, the ambiguities are resolved, thinking and rethinking process results in a more specific formulation of the problem so that it may be a realistic

one in terms of the available data and resources and is also analytically meaningful. All this results in a well defined research problem that is not only meaningful from an operational point of view, but is equally capable of paving the way for the development of working hypotheses and for means of solving the problem itself.

Questions

1. Describe fully the techniques of defining a research problem.
2. What is research problem? Define the main issues which should receive the attention of the researcher in formulating the research problem. Give suitable examples to elucidate your points.

   (*Raj. Uni. EAFM, M. Phil. Exam. 1979*)
3. How do you define a research problem? Give three examples to illustrate your answer.

   (*Raj. Uni. EAFM, M. Phil. Exam. 1978*)
4. What is the necessity of defining a research problem? Explain.
5. Write short notes on:
   (a) Experience survey;
   (b) Pilot survey;
   (c) Components of a research problem;
   (d) Rephrasing the research problem.
6. "The task of defining the research problem often follows a sequential pattern". Explain.
7. "Knowing what data are available often serves to narrow down the problem itself as well as the technique that might be used." Explain the underlying idea in this statement in the context of defining a research problem.
8. Write a comprehensive note on the "Task of defining a research problem".

3

Research Design

MEANING OF RESEARCH DESIGN

The formidable problem that follows the task of defining the research problem is the preparation of the design of the research project, popularly known as the "research design". Decisions regarding what, where, when, how much, by what means concerning an inquiry or a research study constitute a research design. "A research design is the arrangement of conditions for collection and analysis of data in a manner that aims to combine relevance to the research purpose with economy in procedure."[1] In fact, the research design is the conceptual structure within which research is conducted; it constitutes the blueprint for the collection, measurement and analysis of data. As such the design includes an outline of what the researcher will do from writing the hypothesis and its operational implications to the final analysis of data. More explicitly, the desing decisions happen to be in respect of:

(i) What is the study about?
(ii) Why is the study being made?
(iii) Where will the study be carried out?
(iv) What type of data is required?
(v) Where can the required data be found?
(vi) What periods of time will the study include?
(vii) What will be the sample design?
(viii) What techniques of data collection will be used?
(ix) How will the data be analysed?
(x) In what style will the report be prepared?

Keeping in view the above stated design decisions, one may split the overall research design into the following parts:

(a) *the sampling design* which deals with the method of selecting items to be observed for the given study;

[1] Claire Selltiz and others, *Research Methods in Social Sciences*, 1962, p. 50.

(b) *the observational design* which relates to the conditions under which the observations are to be made;

(c) *the statistical design* which concerns with the question of how many items are to be observed and how the information and data gathered are to be analysed; and

(d) *the operational design* which deals with the techniques by which the procedures specified in the sampling, statistical and observational designs can be carried out.

From what has been stated above, we can state the important features of a research design as under:

(i) It is a plan that specifies the sources and types of information relevant to the research problem.

(ii) It is a strategy specifying which approach will be used for gathering and analysing the data.

(iii) It also includes the time and cost budgets since most studies are done under these two constraints.

In brief, research design must, at least, contain—(a) a clear statement of the research problem; (b) procedures and techniques to be used for gathering information; (c) the population to be studied; and (d) methods to be used in processing and analysing data.

NEED FOR RESEARCH DESIGN

Research design is needed because it facilitates the smooth sailing of the various research operations, thereby making research as efficient as possible yielding maximal information with minimal expenditure of effort, time and money. Just as for better, economical and attractive construction of a house, we need a blueprint (or what is commonly called the map of the house) well thought out and prepared by an expert architect, similarly we need a research design or a plan in advance of data collection and analysis for our research project. Research design stands for advance planning of the methods to be adopted for collecting the relevant data and the techniques to be used in their analysis, keeping in view the objective of the research and the availability of staff, time and money. Preparation of the research design should be done with great care as any error in it may upset the entire project. Research design, in fact, has a great bearing on the reliability of the results arrived at and as such constitutes the firm foundation of the entire edifice of the research work.

Even then the need for a well thought out research design is at times not realised by many. The importance which this problem deserves is not given to it. As a result many researches do not serve the purpose for which they are undertaken. In fact, they may even give misleading conclusions. Thoughtlessness in designing the research project may result in rendering the research exercise futile. It is, therefore, imperative that an efficient and appropriate design must be prepared before starting research operations. The design helps the researcher to organize his ideas in a form whereby it will be possible for him to look for flaws and inadequacies. Such a design can even be given to others for their comments and critical evaluation. In the absence of such a course of action, it will be difficult for the critic to provide a comprehensive review of the proposed study.

FEATURES OF A GOOD DESIGN

A good design is often characterised by adjectives like flexible, appropriate, efficient, economical and so on. Generally, the design which minimises bias and maximises the reliability of the data collected and analysed is considered a good design. The design which gives the smallest experimental error is supposed to be the best design in many investigations. Similarly, a design which yields maximal information and provides an opportunity for considering many different aspects of a problem is considered most appropriate and efficient design in respect of many research problems. Thus, the question of good design is related to the purpose or objective of the research problem and also with the nature of the problem to be studied. A design may be quite suitable in one case, but may be found wanting in one respect or the other in the context of some other research problem. One single design cannot serve the purpose of all types of research problems.

A research design appropriate for a particular research problem, usually involves the consideration of the following factors:

(i) the means of obtaining information;
(ii) the availability and skills of the researcher and his staff, if any;
(iii) the objective of the problem to be studied;
(iv) the nature of the problem to be studied; and
(v) the availability of time and money for the research work.

If the research study happens to be an exploratory or a formulative one, wherein the major emphasis is on discovery of ideas and insights, the research design most appropriate must be flexible enough to permit the consideration of many different aspects of a phenomenon. But when the purpose of a study is accurate description of a situation or of an association between variables (or in what are called the descriptive studies), accuracy becomes a major consideration and a research design which minimises bias and maximises the reliability of the evidence collected is considered a good design. Studies involving the testing of a hypothesis of a causal relationship between variables require a design which will permit inferences about causality in addition to the minimisation of bias and maximisation of reliability. But in practice it is the most difficult task to put a particular study in a particular group, for a given research may have in it elements of two or more of the functions of different studies. It is only on the basis of its primary function that a study can be categorised either as an exploratory or descriptive or hypothesis-testing study and accordingly the choice of a research design may be made in case of a particular study. Besides, the availability of time, money, skills of the research staff and the means of obtaining the information must be given due weightage while working out the relevant details of the research design such as experimental design, survey design, sample design and the like.

IMPORTANT CONCEPTS RELATING TO RESEARCH DESIGN

Before describing the different research designs, it will be appropriate to explain the various concepts relating to designs so that these may be better and easily understood.

1. Dependent and independent variables: A concept which can take on different quantitative values is called a variable. As such the concepts like weight, height, income are all examples of variables. Qualitative phenomena (or the attributes) are also quantified on the basis of the presence

or absence of the concerning attribute(s). Phenomena which can take on quantitatively different values even in decimal points are called 'continuous variables'.* But all variables are not continuous. If they can only be expressed in integer values, they are non-continuous variables or in statistical language 'discrete variables'.** Age is an example of continuous variable, but the number of children is an example of non-continuous variable. If one variable depends upon or is a consequence of the other variable, it is termed as a dependent variable, and the variable that is antecedent to the dependent variable is termed as an independent variable. For instance, if we say that height depends upon age, then height is a dependent variable and age is an independent variable. Further, if in addition to being dependent upon age, height also depends upon the individual's sex, then height is a dependent variable and age and sex are independent variables. Similarly, readymade films and lectures are examples of independent variables, whereas behavioural changes, occurring as a result of the environmental manipulations, are examples of dependent variables.

2. **Extraneous variable:** Independent variables that are not related to the purpose of the study, but may affect the dependent variable are termed as extraneous variables. Suppose the researcher wants to test the hypothesis that there is a relationship between children's gains in social studies achievement and their self-concepts. In this case self-concept is an independent variable and social studies achievement is a dependent variable. Intelligence may as well affect the social studies achievement, but since it is not related to the purpose of the study undertaken by the researcher, it will be termed as an extraneous variable. Whatever effect is noticed on dependent variable as a result of extraneous variable(s) is technically described as an 'experimental error'. A study must always be so designed that *the effect upon the dependent variable is attributed entirely to the independent variable(s), and not to some extraneous variable or variables.*

3. **Control:** One important characteristic of a good research design is to minimise the influence or effect of extraneous variable(s). The technical term 'control' is used when we design the study minimising the effects of extraneous independent variables. In experimental researches, the term 'control' is used to refer to restrain experimental conditions.

4. **Confounded relationship:** When the dependent variable is not free from the influence of extraneous variable(s), the relationship between the dependent and independent variables is said to be confounded by an extraneous variable(s).

5. **Research hypothesis:** When a prediction or a hypothesised relationship is to be tested by scientific methods, it is termed as research hypothesis. The research hypothesis is a predictive statement that relates an independent variable to a dependent variable. Usually a research hypothesis must contain, at least, one independent and one dependent variable. Predictive statements which are not to be objectively verified or the relationships that are assumed but not to be tested, are not termed research hypotheses.

6. **Experimental and non-experimental hypothesis-testing research:** When the purpose of research is to test a research hypothesis, it is termed as hypothesis-testing research. It can be of the experimental design or of the non-experimental design. Research in which the independent variable is manipulated is termed 'experimental hypothesis-testing research' and a research in which an independent variable is not manipulated is called 'non-experimental hypothesis-testing research'. For instance, suppose a researcher wants to study whether intelligence affects reading ability for a group

* A continuous variable is that which can assume any numerical value within a specific range.

** A variable for which the individual values fall on the scale only with distinct gaps is called a discrete variable.

of students and for this purpose he randomly selects 50 students and tests their intelligence and reading ability by calculating the coefficient of correlation between the two sets of scores. This is an example of non-experimental hypothesis-testing research because herein the independent variable, intelligence, is not manipulated. But now suppose that our researcher randomly selects 50 students from a group of students who are to take a course in statistics and then divides them into two groups by randomly assigning 25 to Group A, the usual studies programme, and 25 to Group B, the special studies programme. At the end of the course, he administers a test to each group in order to judge the effectiveness of the training programme on the student's performance-level. This is an example of experimental hypothesis-testing research because in this case the independent variable, viz., the type of training programme, is manipulated.

7. Experimental and control groups: In an experimental hypothesis-testing research when a group is exposed to usual conditions, it is termed a 'control group', but when the group is exposed to some novel or special condition, it is termed an 'experimental group'. In the above illustration, the Group A can be called a control group and the Group B an experimental group. If both groups A and B are exposed to special studies programmes, then both groups would be termed 'experimental groups.' It is possible to design studies which include only experimental groups or studies which include both experimental and control groups.

8. Treatments: The different conditions under which experimental and control groups are put are usually referred to as 'treatments'. In the illustration taken above, the two treatments are the usual studies programme and the special studies programme. Similarly, if we want to determine through an experiment the comparative impact of three varieties of fertilizers on the yield of wheat, in that case the three varieties of fertilizers will be treated as three treatments.

9. Experiment: The process of examining the truth of a statistical hypothesis, relating to some research problem, is known as an experiment. For example, we can conduct an experiment to examine the usefulness of a certain newly developed drug. Experiments can be of two types viz., absolute experiment and comparative experiment. If we want to determine the impact of a fertilizer on the yield of a crop, it is a case of absolute experiment; but if we want to determine the impact of one fertilizer as compared to the impact of some other fertilizer, our experiment then will be termed as a comparative experiment. Often, we undertake comparative experiments when we talk of designs of experiments.

10. Experimental unit(s): The pre-determined plots or the blocks, where different treatments are used, are known as experimental units. Such experimental units must be selected (defined) very carefully.

DIFFERENT RESEARCH DESIGNS

Different research designs can be conveniently described if we categorize them as: (1) research design in case of exploratory research studies; (2) research design in case of descriptive and diagnostic research studies, and (3) research design in case of hypothesis-testing research studies.

We take up each category separately.

1. Research design in case of exploratory research studies: Exploratory research studies are also termed as formulative research studies. The main purpose of such studies is that of formulating a problem for more precise investigation or of developing the working hypotheses from an operational

point of view. The major emphasis in such studies is on the discovery of ideas and insights. As such the research design appropriate for such studies must be flexible enough to provide opportunity for considering different aspects of a problem under study. Inbuilt flexibility in research design is needed because the research problem, broadly defined initially, is transformed into one with more precise meaning in exploratory studies, which fact may necessitate changes in the research procedure for gathering relevant data. Generally, the following three methods in the context of research design for such studies are talked about: (a) the survey of concerning literature; (b) the experience survey and (c) the analysis of 'insight-stimulating' examples.

The survey of concerning literature happens to be the most simple and fruitful method of formulating precisely the research problem or developing hypothesis. Hypotheses stated by earlier workers may be reviewed and their usefulness be evaluated as a basis for further research. It may also be considered whether the already stated hypotheses suggest new hypothesis. In this way the researcher should review and build upon the work already done by others, but in cases where hypotheses have not yet been formulated, his task is to review the available material for deriving the relevant hypotheses from it.

Besides, the bibliographical survey of studies, already made in one's area of interest may as well as made by the researcher for precisely formulating the problem. He should also make an attempt to apply concepts and theories developed in different research contexts to the area in which he is himself working. Sometimes the works of creative writers also provide a fertile ground for hypothesis-formulation and as such may be looked into by the researcher.

Experience survey means the survey of people who have had practical experience with the problem to be studied. The object of such a survey is to obtain insight into the relationships between variables and new ideas relating to the research problem. For such a survey people who are competent and can contribute new ideas may be carefully selected as respondents to ensure a representation of different types of experience. The respondents so selected may then be interviewed by the investigator. The researcher must prepare an interview schedule for the systematic questioning of informants. But the interview must ensure flexibility in the sense that the respondents should be allowed to raise issues and questions which the investigator has not previously considered. Generally, the experience-collecting interview is likely to be long and may last for few hours. Hence, it is often considered desirable to send a copy of the questions to be discussed to the respondents well in advance. This will also give an opportunity to the respondents for doing some advance thinking over the various issues involved so that, at the time of interview, they may be able to contribute effectively. Thus, an experience survey may enable the researcher to define the problem more concisely and help in the formulation of the research hypothesis. This survey may as well provide information about the practical possibilities for doing different types of research.

Analysis of 'insight-stimulating' examples is also a fruitful method for suggesting hypotheses for research. It is particularly suitable in areas where there is little experience to serve as a guide. This method consists of the intensive study of selected instances of the phenomenon in which one is interested. For this purpose the existing records, if any, may be examined, the unstructured interviewing may take place, or some other approach may be adopted. Attitude of the investigator, the intensity of the study and the ability of the researcher to draw together diverse information into a unified interpretation are the main features which make this method an appropriate procedure for evoking insights.

Now, what sort of examples are to be selected and studied? There is no clear cut answer to it. Experience indicates that for particular problems certain types of instances are more appropriate than others. One can mention few examples of 'insight-stimulating' cases such as the reactions of strangers, the reactions of marginal individuals, the study of individuals who are in transition from one stage to another, the reactions of individuals from different social strata and the like. In general, cases that provide sharp contrasts or have striking features are considered relatively more useful while adopting this method of hypotheses formulation.

Thus, in an exploratory of formulative research study which merely leads to insights or hypotheses, whatever method or research design outlined above is adopted, the only thing essential is that it must continue to remain flexible so that many different facets of a problem may be considered as and when they arise and come to the notice of the researcher.

2. Research design in case of descriptive and diagnostic research studies: Descriptive research studies are those studies which are concerned with describing the characteristics of a particular individual, or of a group, whereas diagnostic research studies determine the frequency with which something occurs or its association with something else. The studies concerning whether certain variables are associated are examples of diagnostic research studies. As against this, studies concerned with specific predictions, with narration of facts and characteristics concerning individual, group or situation are all examples of descriptive research studies. Most of the social research comes under this category. From the point of view of the research design, the descriptive as well as diagnostic studies share common requirements and as such we may group together these two types of research studies. In descriptive as well as in diagnostic studies, the researcher must be able to define clearly, what he wants to measure and must find adequate methods for measuring it along with a clear cut definition of 'population' he wants to study. Since the aim is to obtain complete and accurate information in the said studies, the procedure to be used must be carefully planned. The research design must make enough provision for protection against bias and must maximise reliability, with due concern for the economical completion of the research study. The design in such studies must be rigid and not flexible and must focus attention on the following:

(a) Formulating the objective of the study (what the study is about and why is it being made?)
(b) Designing the methods of data collection (what techniques of gathering data will be adopted?)
(c) Selecting the sample (how much material will be needed?)
(d) Collecting the data (where can the required data be found and with what time period should the data be related?)
(e) Processing and analysing the data.
(f) Reporting the findings.

In a descriptive/diagnostic study the first step is to specify the objectives with sufficient precision to ensure that the data collected are relevant. If this is not done carefully, the study may not provide the desired information.

Then comes the question of selecting the methods by which the data are to be obtained. In other words, techniques for collecting the information must be devised. Several methods (viz., observation, questionnaires, interviewing, examination of records, etc.), with their merits and limitations, are available for the purpose and the researcher may user one or more of these methods which have been discussed in detail in later chapters. While designing data-collection procedure, adequate safeguards against

bias and unreliability must be ensured. Whichever method is selected, questions must be well examined and be made unambiguous; interviewers must be instructed not to express their own opinion; observers must be trained so that they uniformly record a given item of behaviour. It is always desirable to pre-test the data collection instruments before they are finally used for the study purposes. In other words, we can say that "*structured instruments*" are used in such studies.

In most of the descriptive/diagnostic studies the researcher takes out sample(s) and then wishes to make statements about the population on the basis of the sample analysis or analyses. More often than not, sample has to be designed. Different sample designs have been discussed in detail in a separate chapter in this book. Here we may only mention that the problem of designing samples should be tackled in such a fashion that the samples may yield accurate information with a minimum amount of research effort. Usually one or more forms of probability sampling, or what is often described as random sampling, are used.

To obtain data free from errors introduced by those responsible for collecting them, it is necessary to supervise closely the staff of field workers as they collect and record information. Checks may be set up to ensure that the data collecting staff perform their duty honestly and without prejudice. "As data are collected, they should be examined for completeness, comprehensibility, consistency and reliability."[2]

The data collected must be processed and analysed. This includes steps like coding the interview replies, observations, etc.; tabulating the data; and performing several statistical computations. To the extent possible, the processing and analysing procedure should be planned in detail before actual work is started. This will prove economical in the sense that the researcher may avoid unnecessary labour such as preparing tables for which he later finds he has no use or on the other hand, re-doing some tables because he failed to include relevant data. Coding should be done carefully to avoid error in coding and for this purpose the reliability of coders needs to be checked. Similarly, the accuracy of tabulation may be checked by having a sample of the tables re-done. In case of mechanical tabulation the material (i.e., the collected data or information) must be entered on appropriate cards which is usually done by punching holes corresponding to a given code. The accuracy of punching is to be checked and ensured. Finally, statistical computations are needed and as such averages, percentages and various coefficients must be worked out. Probability and sampling analysis may as well be used. The appropriate statistical operations, along with the use of appropriate tests of significance should be carried out to safeguard the drawing of conclusions concerning the study.

Last of all comes the question of reporting the findings. This is the task of communicating the findings to others and the researcher must do it in an efficient manner. The layout of the report needs to be well planned so that all things relating to the research study may be well presented in simple and effective style.

Thus, the research design in case of descriptive/diagnostic studies is a comparative design throwing light on all points narrated above and must be prepared keeping in view the objective(s) of the study and the resources available. However, it must ensure the minimisation of bias and maximisation of reliability of the evidence collected. The said design can be appropriately referred to as a *survey design* since it takes into account all the steps involved in a survey concerning a phenomenon to be studied.

[2] Claire Selltiz *et al.*, *op. cit.*, p. 74.

The difference between research designs in respect of the above two types of research studies can be conveniently summarised in tabular form as under:

Table 3.1

Research Design	*Type of study*	
	Exploratory of Formulative	*Descriptive/Diagnostic*
Overall design	Flexible design (design must provide opportunity for considering different aspects of the problem)	Rigid design (design must make enough provision for protection against bias and must maximise reliability)
(i) Sampling design	Non-probability sampling design (purposive or judgement sampling)	Probability sampling design (random sampling)
(ii) Statistical design	No pre-planned design for analysis	Pre-planned design for analysis
(iii) Observational design	Unstructured instruments for collection of data	Structured or well thought out instruments for collection of data
(iv) Operational design	No fixed decisions about the operational procedures	Advanced decisions about operational procedures.

3. Research design in case of hypothesis-testing research studies: Hypothesis-testing research studies (generally known as experimental studies) are those where the researcher tests the hypotheses of causal relationships between variables. Such studies require procedures that will not only reduce bias and increase reliability, but will permit drawing inferences about causality. Usually experiments meet this requirement. Hence, when we talk of research design in such studies, we often mean the design of experiments.

Professor R.A. Fisher's name is associated with experimental designs. Beginning of such designs was made by him when he was working at Rothamsted Experimental Station (Centre for Agricultural Research in England). As such the study of experimental designs has its origin in agricultural research. Professor Fisher found that by dividing agricultural fields or plots into different blocks and then by conducting experiments in each of these blocks, whatever information is collected and inferences drawn from them, happens to be more reliable. This fact inspired him to develop certain experimental designs for testing hypotheses concerning scientific investigations. Today, the experimental designs are being used in researches relating to phenomena of several disciplines. Since experimental designs originated in the context of agricultural operations, we still use, though in a technical sense, several terms of agriculture (such as treatment, yield, plot, block etc.) in experimental designs.

BASIC PRINCIPLES OF EXPERIMENTAL DESIGNS

Professor Fisher has enumerated three principles of experimental designs: (1) the Principle of Replication; (2) the Principle of Randomization; and the (3) Principle of Local Control.

According to the *Principle of Replication*, the experiment should be repeated more than once. Thus, each treatment is applied in many experimental units instead of one. By doing so the statistical accuracy of the experiments is increased. For example, suppose we are to examine the effect of two varieties of rice. For this purpose we may divide the field into two parts and grow one variety in one part and the other variety in the other part. We can then compare the yield of the two parts and draw conclusion on that basis. But if we are to apply the principle of replication to this experiment, then we first divide the field into several parts, grow one variety in half of these parts and the other variety in the remaining parts. We can then collect the data of yield of the two varieties and draw conclusion by comparing the same. The result so obtained will be more reliable in comparison to the conclusion we draw without applying the principle of replication. The entire experiment can even be repeated several times for better results. Conceptually replication does not present any difficulty, but computationally it does. For example, if an experiment requiring a two-way analysis of variance is replicated, it will then require a three-way analysis of variance since replication itself may be a source of variation in the data. However, it should be remembered that replication is introduced in order to increase the precision of a study; that is to say, to increase the accuracy with which the main effects and interactions can be estimated.

The *Principle of Randomization* provides protection, when we conduct an experiment, against the effect of extraneous factors by randomization. In other words, this principle indicates that we should design or plan the experiment in such a way that the variations caused by extraneous factors can all be combined under the general heading of "chance." For instance, if we grow one variety of rice, say, in the first half of the parts of a field and the other variety is grown in the other half, then it is just possible that the soil fertility may be different in the first half in comparison to the other half. If this is so, our results would not be realistic. In such a situation, we may assign the variety of rice to be grown in different parts of the field on the basis of some random sampling technique i.e., we may apply randomization principle and protect ourselves against the effects of the extraneous factors (soil fertility differences in the given case). As such, through the application of the principle of randomization, we can have a better estimate of the experimental error.

The *Principle of Local Control* is another important principle of experimental designs. Under it the extraneous factor, the known source of variability, is made to vary deliberately over as wide a range as necessary and this needs to be done in such a way that the variability it causes can be measured and hence eliminated from the experimental error. This means that we should plan the experiment in a manner that we can perform a two-way analysis of variance, in which the total variability of the data is divided into three components attributed to treatments (varieties of rice in our case), the extraneous factor (soil fertility in our case) and experimental error.* In other words, according to the principle of local control, we first divide the field into several homogeneous parts, known as blocks, and then each such block is divided into parts equal to the number of treatments. Then the treatments are randomly assigned to these parts of a block. Dividing the field into several homogenous parts is known as 'blocking'. In general, blocks are the levels at which we hold an extraneous factor fixed, so that we can measure its contribution to the total variability of the data by means of a two-way analysis of variance. In brief, through the principle of local control we can eliminate the variability due to extraneous factor(s) from the experimental error.

* See Chapter Analysis of Variance for details.

Important Experimental Designs

Experimental design refers to the framework or structure of an experiment and as such there are several experimental designs. We can classify experimental designs into two broad categories, viz., informal experimental designs and formal experimental designs. Informal experimental designs are those designs that normally use a less sophisticated form of analysis based on differences in magnitudes, whereas formal experimental designs offer relatively more control and use precise statistical procedures for analysis. Important experiment designs are as follows:

(*a*) Informal experimental designs:
  (i) Before-and-after without control design.
  (ii) After-only with control design.
  (iii) Before-and-after with control design.

(*b*) Formal experimental designs:
  (i) Completely randomized design (C.R. Design).
  (ii) Randomized block design (R.B. Design).
  (iii) Latin square design (L.S. Design).
  (iv) Factorial designs.

We may briefly deal with each of the above stated informal as well as formal experimental designs.

1. Before-and-after without control design: In such a design a single test group or area is selected and the dependent variable is measured before the introduction of the treatment. The treatment is then introduced and the dependent variable is measured again after the treatment has been introduced. The effect of the treatment would be equal to the level of the phenomenon after the treatment minus the level of the phenomenon before the treatment. The design can be represented thus:

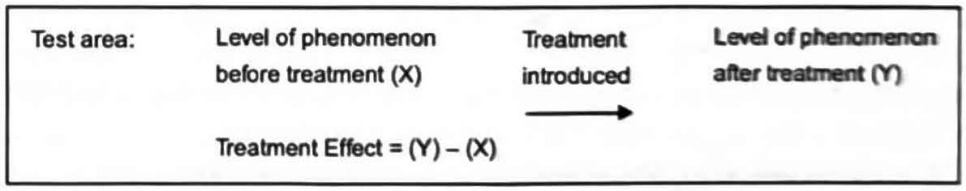

Test area:
Level of phenomenon before treatment (X)
Treatment introduced
Level of phenomenon after treatment (Y)
Treatment Effect = (Y) – (X)

Fig. 3.1

The main difficulty of such a design is that with the passage of time considerable extraneous variations may be there in its treatment effect.

2. After-only with control design: In this design two groups or areas (test area and control area) are selected and the treatment is introduced into the test area only. The dependent variable is then measured in both the areas at the same time. Treatment impact is assessed by subtracting the value of the dependent variable in the control area from its value in the test area. This can be exhibited in the following form:

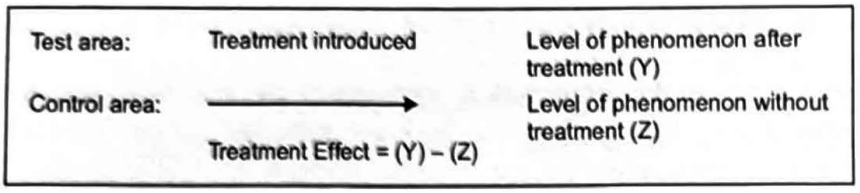

Test area:
Treatment introduced
Level of phenomenon after treatment (Y)
Control area:
Level of phenomenon without treatment (Z)
Treatment Effect = (Y) – (Z)

Fig. 3.2

The basic assumption in such a design is that the two areas are identical with respect to their behaviour towards the phenomenon considered. If this assumption is not true, there is the possibility of extraneous variation entering into the treatment effect. However, data can be collected in such a design without the introduction of problems with the passage of time. In this respect the design is superior to before-and-after without control design.

3. **Before-and-after with control design:** In this design two areas are selected and the dependent variable is measured in both the areas for an identical time-period before the treatment. The treatment is then introduced into the test area only, and the dependent variable is measured in both for an identical time-period after the introduction of the treatment. The treatment effect is determined by subtracting the change in the dependent variable in the control area from the change in the dependent variable in test area. This design can be shown in this way:

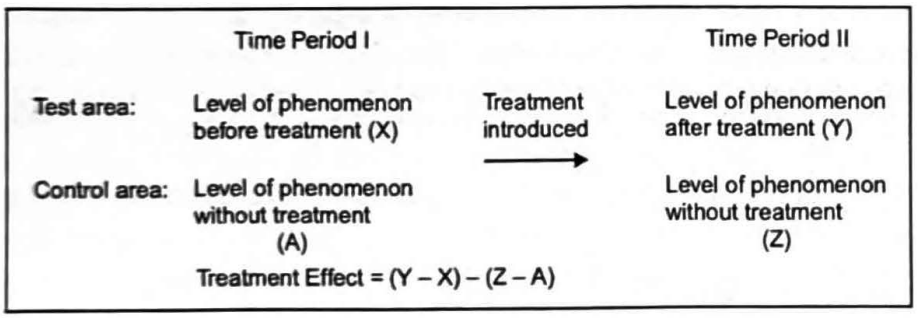

Time Period I
Time Period II
Test area:
Level of phenomenon before treatment (X)
Treatment introduced
Level of phenomenon after treatment (Y)
Control area:
Level of phenomenon without treatment (A)
Level of phenomenon without treatment (Z)
Treatment Effect = (Y – X) – (Z – A)

Fig. 3.3

This design is superior to the above two designs for the simple reason that it avoids extraneous variation resulting both from the passage of time and from non-comparability of the test and control areas. But at times, due to lack of historical data, time or a comparable control area, we should prefer to select one of the first two informal designs stated above.

4. **Completely randomized design (C.R. design):** Involves only two principles viz., the principle of replication and the principle of randomization of experimental designs. It is the simplest possible design and its procedure of analysis is also easier. The essential characteristic of the design is that subjects are randomly assigned to experimental treatments (or vice-versa). For instance, if we have 10 subjects and if we wish to test 5 under treatment A and 5 under treatment B, the randomization process gives every possible group of 5 subjects selected from a set of 10 an equal opportunity of being assigned to treatment A and treatment B. One-way analysis of variance (or one-way ANOVA)* is used to analyse such a design. Even unequal replications can also work in this design. It provides maximum number of degrees of freedom to the error. Such a design is generally used when experimental areas happen to be homogeneous. Technically, when all the variations due to uncontrolled

* See Chapter 11 for one-way ANOVA technique.

extraneous factors are included under the heading of chance variation, we refer to the design of experiment as C.R. design.

We can present a brief description of the two forms of such a design as given in Fig 3.4.

(i) Two-group simple randomized design: In a two-group simple randomized design, first of all the population is defined and then from the population a sample is selected randomly. Further, requirement of this design is that items, after being selected randomly from the population, be randomly assigned to the experimental and control groups (Such random assignment of items to two groups is technically described as principle of randomization). Thus, this design yields two groups as representatives of the population. In a diagram form this design can be shown in this way:

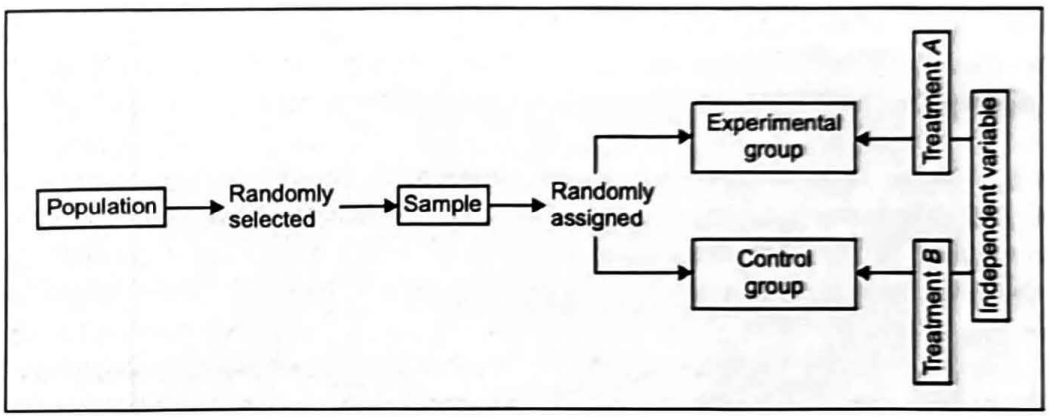


Fig. 3.4: Two-group simple randomized experimental design (in diagram form)

Since in the sample randomized design the elements constituting the sample are randomly drawn from the same population and randomly assigned to the experimental and control groups, it becomes possible to draw conclusions on the basis of samples applicable for the population. The two groups (experimental and control groups) of such a design are given different treatments of the independent variable. This design of experiment is quite common in research studies concerning behavioural sciences. The merit of such a design is that it is simple and randomizes the differences among the sample items. But the limitation of it is that the individual differences among those conducting the treatments are not eliminated, i.e., it does not control the extraneous variable and as such the result of the experiment may not depict a correct picture. This can be illustrated by taking an example. Suppose the researcher wants to compare two groups of students who have been randomly selected and randomly assigned. Two different treatments viz., the usual training and the specialised training are being given to the two groups. The researcher hypothesises greater gains for the group receiving specialised training. To determine this, he tests each group before and after the training, and then compares the amount of gain for the two groups to accept or reject his hypothesis. This is an illustration of the two-groups randomized design, wherein individual differences among students are being randomized. But this does not control the differential effects of the extraneous independent variables (in this case, the individual differences among those conducting the training programme).

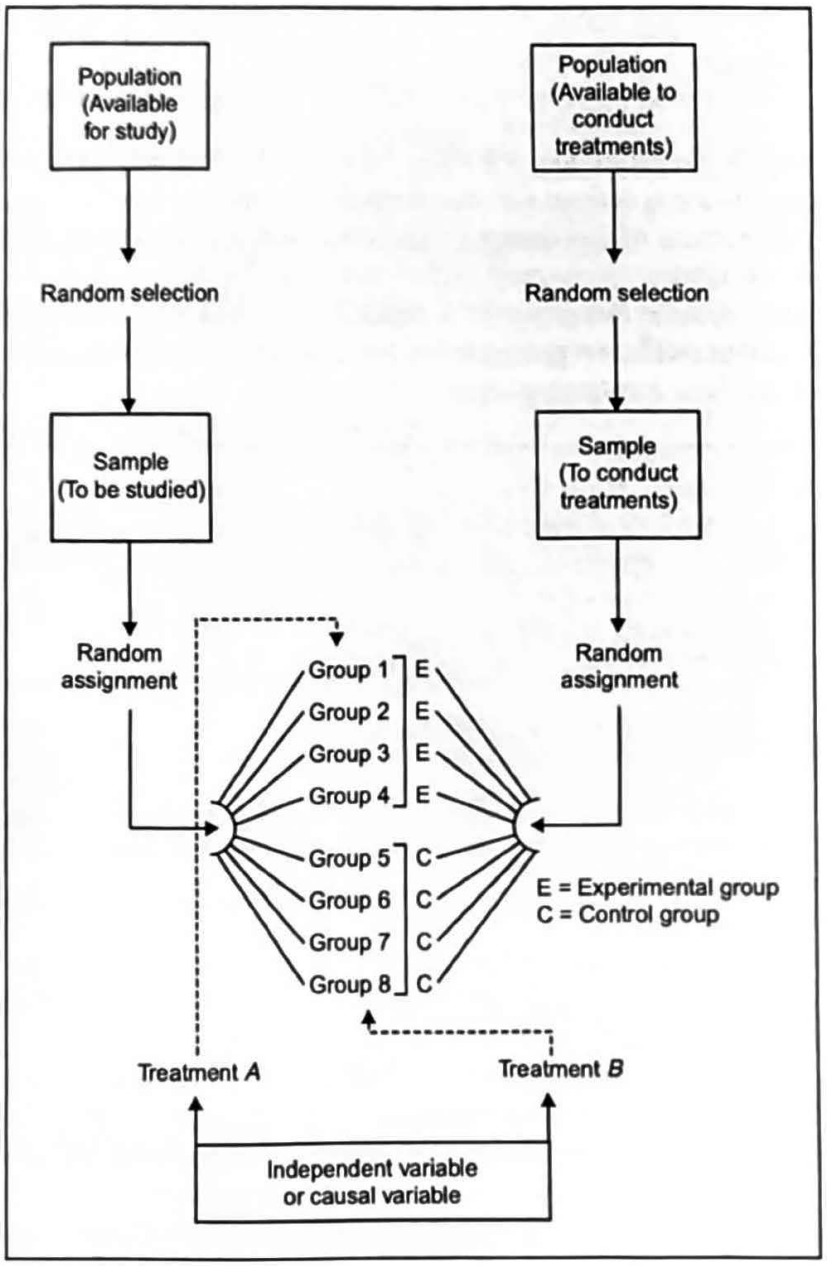


Fig. 3.5: Random replication design (in diagram form)

(ii) **Random replications design:** The limitation of the two-group randomized design is usually eliminated within the random replications design. In the illustration just cited above, the *teacher differences* on the dependent variable were ignored, i.e., the extraneous variable was not controlled. But in a random replications design, the effect of such differences are minimised (or reduced) by providing a number of repetitions for each treatment. Each repetition is technically called a 'replication'. Random replication design serves two purposes viz., it provides controls for the differential effects of the extraneous independent variables and secondly, it randomizes any individual differences among those conducting the treatments. Diagrammatically we can illustrate the random replications design thus: (Fig. 3.5)

From the diagram it is clear that there are two populations in the replication design. The sample is taken randomly from the population available for study and is randomly assigned to, say, four experimental and four control groups. Similarly, sample is taken randomly from the population available to conduct experiments (because of the eight groups eight such individuals be selected) and the eight individuals so selected should be randomly assigned to the eight groups. Generally, equal number of items are put in each group so that the size of the group is not likely to affect the result of the study. Variables relating to both population characteristics are assumed to be randomly distributed among the two groups. Thus, this random replication design is, in fact, an extension of the two-group simple randomized design.

5. **Randomized block design (R.B. design)** is an improvement over the C.R. design. In the R.B. design the principle of local control can be applied along with the other two principles of experimental designs. In the R.B. design, subjects are first divided into groups, known as blocks, such that within each group the subjects are relatively homogeneous in respect to some selected variable. The variable selected for grouping the subjects is one that is believed to be related to the measures to be obtained in respect of the dependent variable. The number of subjects in a given block would be equal to the number of treatments and one subject in each block would be randomly assigned to each treatment. In general, blocks are the levels at which we hold the extraneous factor fixed, so that its contribution to the total variability of data can be measured. The main feature of the R.B. design is that in this each treatment appears the same number of times in each block. The R.B. design is analysed by the two-way analysis of variance (two-way ANOVA)* technique.

Let us illustrate the R.B. design with the help of an example. Suppose four different forms of a standardised test in statistics were given to each of five students (selected one from each of the five I.Q. blocks) and following are the scores which they obtained.

	Very low I.Q.	Low I.Q.	Average I.Q.	High I.Q.	Very high I.Q.
	Student A	Student B	Student C	Student D	Student E
Form 1	82	67	57	71	73
Form 2	90	68	54	70	81
Form 3	86	73	51	69	84
Form 4	93	77	60	65	71

Fig. 3.6

If each student separately randomized the order in which he or she took the four tests (by using random numbers or some similar device), we refer to the design of this experiment as a R.B. design. The purpose of this randomization is to take care of such possible extraneous factors (say as fatigue) or perhaps the experience gained from repeatedly taking the test.

* See Chapter 11 for the two-way ANOVA technique.

6. Latin square design (L.S. design) is an experimental design very frequently used in agricultural research. The conditions under which agricultural investigations are carried out are different from those in other studies for nature plays an important role in agriculture. For instance, an experiment has to be made through which the effects of five different varieties of fertilizers on the yield of a certain crop, say wheat, it to be judged. In such a case the varying fertility of the soil in different blocks in which the experiment has to be performed must be taken into consideration; otherwise the results obtained may not be very dependable because the output happens to be the effect not only of fertilizers, but it may also be the effect of fertility of soil. Similarly, there may be impact of varying seeds on the yield. To overcome such difficulties, the L.S. design is used when there are two major extraneous factors such as the varying soil fertility and varying seeds.

The Latin-square design is one wherein each fertilizer, in our example, appears five times but is used only once in each row and in each column of the design. In other words, the treatments in a L.S. design are so allocated among the plots that no treatment occurs more than once in any one row or any one column. The two blocking factors may be represented through rows and columns (one through rows and the other through columns). The following is a diagrammatic form of such a design in respect of, say, five types of fertilizers, viz., A, B, C, D and E and the two blocking factor viz., the varying soil fertility and the varying seeds:

FERTILITY LEVEL

Seeds differences

	I	II	III	IV	V
X_1	A	B	C	D	E
X_2	B	C	D	E	A
X_3	C	D	E	A	B
X_4	D	E	A	B	C
X_5	E	A	B	C	D

Fig. 3.7

The above diagram clearly shows that in a L.S. design the field is divided into as many blocks as there are varieties of fertilizers and then each block is again divided into as many parts as there are varieties of fertilizers in such a way that each of the fertilizer variety is used in each of the block (whether column-wise or row-wise) only once. The analysis of the L.S. design is very similar to the two-way ANOVA technique.

The merit of this experimental design is that it enables differences in fertility gradients in the field to be eliminated in comparison to the effects of different varieties of fertilizers on the yield of the crop. But this design suffers from one limitation, and it is that although each row and each column represents equally all fertilizer varieties, there may be considerable difference in the row and column means both up and across the field. This, in other words, means that in L.S. design we must assume that there is no interaction between treatments and blocking factors. This defect can, however, be removed by taking the means of rows and columns equal to the field mean by adjusting the results. Another limitation of this design is that it requires number of rows, columns and treatments to be

equal. This reduces the utility of this design. In case of (2 × 2) L.S. design, there are no degrees of freedom available for the mean square error and hence the design cannot be used. If treatments are 10 or more, than each row and each column will be larger in size so that rows and columns may not be homogeneous. This may make the application of the principle of local control ineffective. Therefore, L.S. design of orders (5 × 5) to (9 × 9) are generally used.

7. **Factorial designs:** Factorial designs are used in experiments where the effects of varying more than one factor are to be determined. They are specially important in several economic and social phenomena where usually a large number of factors affect a particular problem. Factorial designs can be of two types: (i) simple factorial designs and (ii) complex factorial designs. We take them separately

(i) *Simple factorial designs:* In case of simple factorial designs, we consider the effects of varying two factors on the dependent variable, but when an experiment is done with more than two factors, we use complex factorial designs. Simple factorial design is also termed as a 'two-factor-factorial design', whereas complex factorial design is known as 'multi-factor-factorial design.' Simple factorial design may either be a 2 × 2 simple factorial design, or it may be, say, 3 × 4 or 5 × 3 or the like type of simple factorial design. We illustrate some simple factorial designs as under:

Illustration 1: (2 × 2 simple factorial design).

A 2 × 2 simple factorial design can graphically be depicted as follows:

2 × 2 SIMPLE FACTORIAL DESIGN

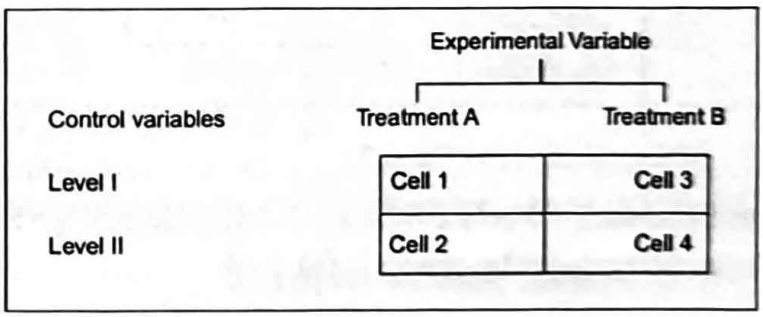


Fig. 3.8

In this design the extraneous variable to be controlled by homogeneity is called the control variable and the independent variable, which is manipulated, is called the experimental variable. Then there are two treatments of the experimental variable and two levels of the control variable. As such there are four cells into which the sample is divided. Each of the four combinations would provide one treatment or experimental condition. Subjects are assigned at random to each treatment in the same manner as in a randomized group design. The means for different cells may be obtained along with the means for different rows and columns. Means of different cells represent the mean scores for the dependent variable and the column means in the given design are termed the main effect for treatments without taking into account any differential effect that is due to the level of the control variable. Similarly, the row means in the said design are termed the main effects for levels without regard to treatment. Thus, through this design we can study the main effects of treatments as well as

the main effects of levels. An additional merit of this design is that one can examine the interaction between treatments and levels, through which one may say whether the treatment and levels are independent of each other or they are not so. The following examples make clear the interaction effect between treatments and levels. The data obtained in case of two (2 × 2) simple factorial studies may be as given in Fig. 3.9.

STUDY I DATA

		Training		
		Treatment A	Treatment B	Row Mean
Control (Intelligence)	Level I (Low)	15.5	23.3	19.4
	Level II (High)	35.8	30.2	33.0
	Column mean	25.6	26.7	

STUDY II DATA

		Training		
		Treatment A	Treatment B	Row Mean
Control (Intelligence)	Level I (Low)	10.4	20.6	15.5
	Level II (High)	30.6	40.4	35.5
	Column mean	20.5	30.5	

Fig. 3.9

All the above figures (the study I data and the study II data) represent the respective means. Graphically, these can be represented as shown in Fig. 3.10.

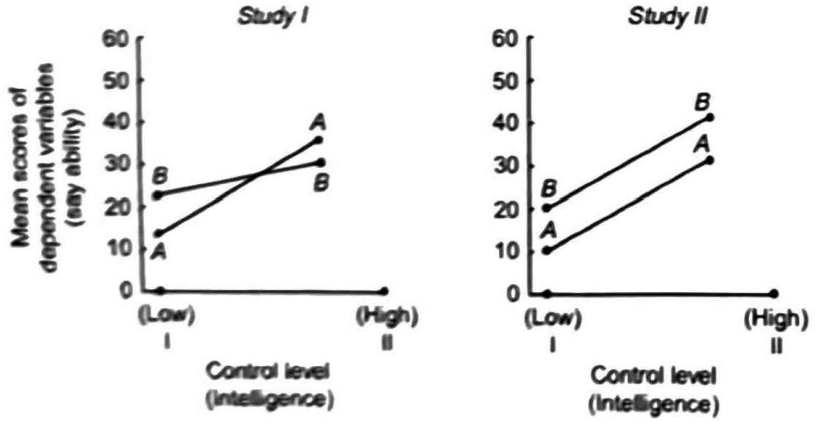


Fig. 3.10

The graph relating to Study I indicates that there is an interaction between the treatment and the level which, in other words, means that the treatment and the level are not independent of each other. The graph relating to Study II shows that there is no interaction effect which means that treatment and level in this study are relatively independent of each other.

The 2 × 2 design need not be restricted in the manner as explained above i.e., having one experimental variable and one control variable, but it may also be of the type having two experimental variables or two control variables. For example, a college teacher compared the effect of the class-size as well as the introduction of the new instruction technique on the learning of research methodology. For this purpose he conducted a study using a 2 × 2 simple factorial design. His design in the graphic form would be as follows:

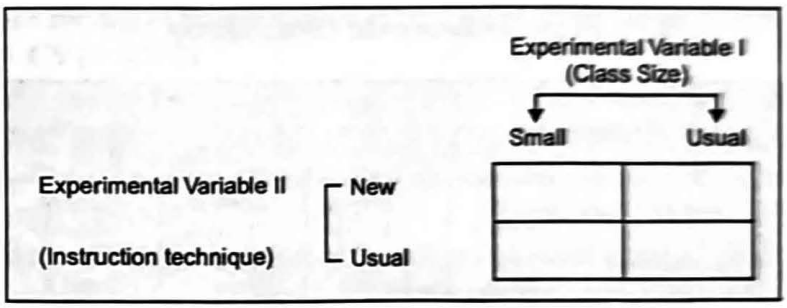


Fig. 3.11

But if the teacher uses a design for comparing males and females and the senior and junior students in the college as they relate to the knowledge of research methodology, in that case we will have a 2 × 2 simple factorial design wherein both the variables are control variables as no manipulation is involved in respect of both the variables.

Illustration 2: (4 × 3 simple factorial design).

The 4 × 3 simple factorial design will usually include four treatments of the experimental variable and three levels of the control variable. Graphically it may take the following form:

4 × 3 SIMPLE FACTORIAL DESIGN

Control Variable	Experimental Variable			
	Treatment A	Treatment B	Treatment C	Treatment D
Level I	Cell 1	Cell 4	Cell 7	Cell 10
Level II	Cell 2	Cell 5	Cell 8	Cell 11
Level III	Cell 3	Cell 6	Cell 9	Cell 12

Fig. 3.12

This model of a simple factorial design includes four treatments viz., A, B, C, and D of the experimental variable and three levels viz., I, II, and III of the control variable and has 12 different cells as shown above. This shows that a 2 × 2 simple factorial design can be generalised to any number of treatments and levels. Accordingly we can name it as such and such (–×–) design. In

such a design the means for the columns provide the researcher with an estimate of the main effects for treatments and the means for rows provide an estimate of the main effects for the levels. Such a design also enables the researcher to determine the interaction between treatments and levels.

(ii) *Complex factorial designs:* Experiments with more than two factors at a time involve the use of complex factorial designs. A design which considers three or more independent variables simultaneously is called a complex factorial design. In case of three factors with one experimental variable having two treatments and two control variables, each one of which having two levels, the design used will be termed 2 × 2 × 2 complex factorial design which will contain a total of eight cells as shown below in Fig. 3.13.

2 × 2 × 2 COMPLEX FACTORIAL DESIGN

		Experimental Variable			
		Treatment A		Treatment B	
		Control Variable 2 Level I	Control Variable 2 Level II	Control Variable 2 Level I	Control Variable 2 Level II
Control Variable 1	Level I	Cell 1	Cell 3	Cell 5	Cell 7
	Level II	Cell 2	Cell 4	Cell 6	Cell 8

Fig. 3.13

In Fig. 3.14 a pictorial presentation is given of the design shown below.

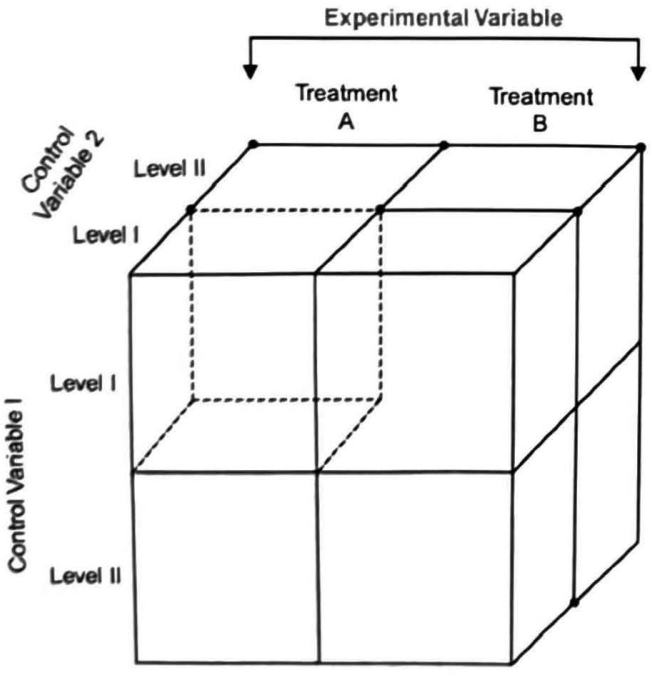


Fig. 3.14

The dotted line cell in the diagram corresponds to Cell 1 of the above stated 2 × 2 × 2 design and is for Treatment A, level I of the control variable 1, and level I of the control variable 2. From this design it is possible to determine the main effects for three variables i.e., one experimental and two control variables. The researcher can also determine the interactions between each possible pair of variables (such interactions are called 'First Order interactions') and interaction between variable taken in triplets (such interactions are called Second Order interactions). In case of a 2 × 2 × 2 design, the further given first order interactions are possible:

Experimental variable with control variable 1 (or EV × CV 1);

Experimental variable with control variable 2 (or EV × CV 2);

Control variable 1 with control variable 2 (or CV1 × CV2);

Three will be one second order interaction as well in the given design (it is between all the three variables i.e., EV × CV1 × CV2).

To determine the main effects for the experimental variable, the researcher must necessarily compare the combined mean of data in cells 1, 2, 3 and 4 for Treatment A with the combined mean of data in cells 5, 6, 7 and 8 for Treatment B. In this way the main effect for experimental variable, independent of control variable 1 and variable 2, is obtained. Similarly, the main effect for control variable 1, independent of experimental variable and control variable 2, is obtained if we compare the combined mean of data in cells 1, 3, 5 and 7 with the combined mean of data in cells 2, 4, 6 and 8 of our 2 × 2 × 2 factorial design. On similar lines, one can determine the main effect for the control variable 2 independent of experimental variable and control variable 1, if the combined mean of data in cells 1, 2, 5 and 6 are compared with the combined mean of data in cells 3, 4, 7 and 8.

To obtain the first order interaction, say, for EV × CV1 in the above stated design, the researcher must necessarily ignore control variable 2 for which purpose he may develop 2 × 2 design from the 2 × 2 × 2 design by combining the data of the relevant cells of the latter design as shown in Fig. 3.15.

		Experimental Variables	
		Treatment A	Treatment B
Control Variable 1	Level I	Cells 1, 3	Cells 5, 7
	Level II	Cells 2, 4	Cells 6, 8

Fig. 3.15

Similarly, the researcher can determine other first order interactions. The analysis of the first order interaction, in the manner described above, is essentially a sample factorial analysis as only two variables are considered at a time and the remaining one is ignored. But the analysis of the second order interaction would not ignore one of the three independent variables in case of a 2 × 2 × 2 design. The analysis would be termed as a complex factorial analysis.

It may, however, be remembered that the complex factorial design need not necessarily be of 2 × 2 × 2 type design, but can be generalised to any number and combination of experimental and control independent variables. Of course, the greater the number of independent variables included in a complex factorial design, the higher the order of the interaction analysis possible. But the overall task goes on becoming more and more complicated with the inclusion of more and more independent variables in our design.

Factorial designs are used mainly because of the two advantages. (i) They provide equivalent accuracy (as happens in the case of experiments with only one factor) with less labour and as such are a source of economy. Using factorial designs, we can determine the main effects of two (in simple factorial design) or more (in case of complex factorial design) factors (or variables) in one single experiment. (ii) They permit various other comparisons of interest. For example, they give information about such effects which cannot be obtained by treating one single factor at a time. The determination of interaction effects is possible in case of factorial designs.

CONCLUSION

There are several research designs and the researcher must decide in advance of collection and analysis of data as to which design would prove to be more appropriate for his research project. He must give due weight to various points such as the type of universe and its nature, the objective of his study, the resource list or the sampling frame, desired standard of accuracy and the like when taking a decision in respect of the design for his research project.

Questions

1. Explain the meaning and significance of a Research design.
2. Explain the meaning of the following in context of Research design.
   (a) Extraneous variables;
   (b) Confounded relationship;
   (c) Research hypothesis;
   (d) Experimental and Control groups;
   (e) Treatments.
3. Describe some of the important research designs used in experimental hypothesis-testing research study.
4. "Research design in exploratory studies must be flexible but in descriptive studies, it must minimise bias and maximise reliability." Discuss.
5. Give your understanding of a good research design. Is single research design suitable in all research studies? If not, why?
6. Explain and illustrate the following research designs:
   (a) Two group simple randomized design;
   (b) Latin square design;
   (c) Random replications design;
   (d) Simple factorial design;
   (e) Informal experimental designs.
7. Write a short note on 'Experience Survey' explaining fully its utility in exploratory research studies.
8. What is research design? Discuss the basis of stratification to be employed in sampling public opinion on inflation.

(Raj. Uni. EAFM M. Phil, Exam. 1978)

Appendix

Developing a Research Plan*

After identifying and defining the problem as also accomplishing the relating task, researcher must arrange his ideas in order and write them in the form of an experimental plan or what can be described as 'Research Plan'. This is essential specially for new researcher because of the following:

(a) It helps him to organize his ideas in a form whereby it will be possible for him to look for flaws and inadequacies, if any.

(b) It provides an inventory of what must be done and which materials have to be collected as a preliminary step.

(c) It is a document that can be given to others for comment.

Research plan must contain the following items.

1. Research objective should be clearly stated in a line or two which tells exactly what it is that the researcher expects to do.
2. The problem to be studied by researcher must be explicitly stated so that one may know what information is to be obtained for solving the problem.
3. Each major concept which researcher wants to measure should be defined in operational terms in context of the research project.
4. The plan should contain the method to be used in solving the problem. An overall description of the approach to be adopted is usually given and assumptions, if any, of the concerning method to be used are clearly mentioned in the research plan.
5. The plan must also state the details of the techniques to be adopted. For instance, if interview method is to be used, an account of the nature of the contemplated interview procedure should be given. Similarly, if tests are to be given, the conditions under which they are to be administered should be specified along with the nature of instruments to be used. If public records are to be consulted as sources of data, the fact should be recorded in the research plan. Procedure for quantifying data should also be written out in all details.

* Based on the matter given in the following two books:

(i) Robert M.W. Travers, *An Introduction to Educational Research*, p. 82–84.

(ii) C. William Emory, *Business Research Methods*, p. 415–416.

6. A clear mention of the population to be studied should be made. If the study happens to be sample based, the research plan should state the sampling plan i.e., how the sample is to be identified. The method of identifying the sample should be such that generalisation from the sample to the original population is feasible.
7. The plan must also contain the methods to be used in processing the data. Statistical and other methods to be used must be indicated in the plan. Such methods should not be left until the data have been collected. This part of the plan may be reviewed by experts in the field, for they can often suggest changes that result in substantial saving of time and effort.
8. Results of pilot test, if any, should be reported. Time and cost budgets for the research project should also be prepared and laid down in the plan itself.

4

Sampling Design

CENSUS AND SAMPLE SURVEY

All items in any field of inquiry constitute a 'Universe' or 'Population.' A complete enumeration of all items in the 'population' is known as a census inquiry. It can be presumed that in such an inquiry, when all items are covered, no element of chance is left and highest accuracy is obtained. But in practice this may not be true. Even the slightest element of bias in such an inquiry will get larger and larger as the number of observation increases. Moreover, there is no way of checking the element of bias or its extent except through a resurvey or use of sample checks. Besides, this type of inquiry involves a great deal of time, money and energy. Therefore, when the field of inquiry is large, this method becomes difficult to adopt because of the resources involved. At times, this method is practically beyond the reach of ordinary researchers. Perhaps, government is the only institution which can get the complete enumeration carried out. Even the government adopts this in very rare cases such as population census conducted once in a decade. Further, many a time it is not possible to examine every item in the population, and sometimes it is possible to obtain sufficiently accurate results by studying only a part of total population. In such cases there is no utility of census surveys.

However, it needs to be emphasised that when the universe is a small one, it is no use resorting to a sample survey. When field studies are undertaken in practical life, considerations of time and cost almost invariably lead to a selection of respondents i.e., selection of only a few items. The respondents selected should be as representative of the total population as possible in order to produce a miniature cross-section. The selected respondents constitute what is technically called a 'sample' and the selection process is called 'sampling technique.' The survey so conducted is known as 'sample survey'. Algebraically, let the population size be N and if a part of size n (which is $< N$) of this population is selected according to some rule for studying some characteristic of the population, the group consisting of these n units is known as 'sample'. Researcher must prepare a sample design for his study i.e., he must plan how a sample should be selected and of what size such a sample would be.

IMPLICATIONS OF A SAMPLE DESIGN

A sample design is a definite plan for obtaining a sample from a given population. It refers to the technique or the procedure the researcher would adopt in selecting items for the sample. Sample

design may as well lay down the number of items to be included in the sample i.e., the size of the sample. Sample design is determined before data are collected. There are many sample designs from which a researcher can choose. Some designs are relatively more precise and easier to apply than others. Researcher must select/prepare a sample design which should be reliable and appropriate for his research study.

STEPS IN SAMPLE DESIGN

While developing a sampling design, the researcher must pay attention to the following points:

(i) Type of universe: The first step in developing any sample design is to clearly define the set of objects, technically called the Universe, to be studied. The universe can be finite or infinite. In finite universe the number of items is certain, but in case of an infinite universe the number of items is infinite, i.e., we cannot have any idea about the total number of items. The population of a city, the number of workers in a factory and the like are examples of finite universes, whereas the number of stars in the sky, listeners of a specific radio programme, throwing of a dice etc. are examples of infinite universes.

(ii) Sampling unit: A decision has to be taken concerning a sampling unit before selecting sample. Sampling unit may be a geographical one such as state, district, village, etc., or a construction unit such as house, flat, etc., or it may be a social unit such as family, club, school, etc., or it may be an individual. The researcher will have to decide one or more of such units that he has to select for his study.

(iii) Source list: It is also known as 'sampling frame' from which sample is to be drawn. It contains the names of all items of a universe (in case of finite universe only). If source list is not available, researcher has to prepare it. Such a list should be comprehensive, correct, reliable and appropriate. It is extremely important for the source list to be as representative of the population as possible.

(iv) Size of sample: This refers to the number of items to be selected from the universe to constitute a sample. This a major problem before a researcher. The size of sample should neither be excessively large, nor too small. It should be optimum. An optimum sample is one which fulfills the requirements of efficiency, representativeness, reliability and flexibility. While deciding the size of sample, researcher must determine the desired precision as also an acceptable confidence level for the estimate. The size of population variance needs to be considered as in case of larger variance usually a bigger sample is needed. The size of population must be kept in view for this also limits the sample size. The parameters of interest in a research study must be kept in view, while deciding the size of the sample. Costs too dictate the size of sample that we can draw. As such, budgetary constraint must invariably be taken into consideration when we decide the sample size.

(v) Parameters of interest: In determining the sample design, one must consider the question of the specific population parameters which are of interest. For instance, we may be interested in estimating the proportion of persons with some characteristic in the population, or we may be interested in knowing some average or the other measure concerning the population. There may also be important sub-groups in the population about whom we

would like to make estimates. All this has a strong impact upon the sample design we would accept.

(vi) **Budgetary constraint:** Cost considerations, from practical point of view, have a major impact upon decisions relating to not only the size of the sample but also to the type of sample. This fact can even lead to the use of a non-probability sample.

(vii) **Sampling procedure:** Finally, the researcher must decide the type of sample he will use i.e., he must decide about the technique to be used in selecting the items for the sample. In fact, this technique or procedure stands for the sample design itself. There are several sample designs (explained in the pages that follow) out of which the researcher must choose one for his study. Obviously, he must select that design which, for a given sample size and for a given cost, has a smaller sampling error.

CRITERIA OF SELECTING A SAMPLING PROCEDURE

In this context one must remember that two costs are involved in a sampling analysis viz., the cost of collecting the data and the cost of an incorrect inference resulting from the data. Researcher must keep in view the two causes of incorrect inferences viz., systematic bias and sampling error. A *systematic bias* results from errors in the sampling procedures, and it cannot be reduced or eliminated by increasing the sample size. At best the causes responsible for these errors can be detected and corrected. Usually a systematic bias is the result of one or more of the following factors:

1. **Inappropriate sampling frame:** If the sampling frame is inappropriate i.e., a biased representation of the universe, it will result in a systematic bias.

2. **Defective measuring device:** If the measuring device is constantly in error, it will result in systematic bias. In survey work, systematic bias can result if the questionnaire or the interviewer is biased. Similarly, if the physical measuring device is defective there will be systematic bias in the data collected through such a measuring device.

3. **Non-respondents:** If we are unable to sample all the individuals initially included in the sample, there may arise a systematic bias. The reason is that in such a situation the likelihood of establishing contact or receiving a response from an individual is often correlated with the measure of what is to be estimated.

4. **Indeterminancy principle:** Sometimes we find that individuals act differently when kept under observation than what they do when kept in non-observed situations. For instance, if workers are aware that somebody is observing them in course of a work study on the basis of which the average length of time to complete a task will be determined and accordingly the quota will be set for piece work, they generally tend to work slowly in comparison to the speed with which they work if kept unobserved. Thus, the indeterminancy principle may also be a cause of a systematic bias.

5. **Natural bias in the reporting of data:** Natural bias of respondents in the reporting of data is often the cause of a systematic bias in many inquiries. There is usually a downward bias in the income data collected by government taxation department, whereas we find an upward bias in the income data collected by some social organisation. People in general understate their incomes if asked about it for tax purposes, but they overstate the same if asked for social status or their affluence. Generally in psychological surveys, people tend to give what they think is the 'correct' answer rather than revealing their true feelings.

Sampling errors are the random variations in the sample estimates around the true population parameters. Since they occur randomly and are equally likely to be in either direction, their nature happens to be of compensatory type and the expected value of such errors happens to be equal to zero. Sampling error decreases with the increase in the size of the sample, and it happens to be of a smaller magnitude in case of homogeneous population.

Sampling error can be measured for a given sample design and size. The measurement of sampling error is usually called the 'precision of the sampling plan'. If we increase the sample size, the precision can be improved. But increasing the size of the sample has its own limitations viz., a large sized sample increases the cost of collecting data and also enhances the systematic bias. Thus the effective way to increase precision is usually to select a better sampling design which has a smaller sampling error for a given sample size at a given cost. In practice, however, people prefer a less precise design because it is easier to adopt the same and also because of the fact that systematic bias can be controlled in a better way in such a design.

In brief, *while selecting a sampling procedure, researcher must ensure that the procedure causes a relatively small sampling error and helps to control the systematic bias in a better way.*

CHARACTERISTICS OF A GOOD SAMPLE DESIGN

From what has been stated above, we can list down the characteristics of a good sample design as under:

(a) Sample design must result in a truly representative sample.
(b) Sample design must be such which results in a small sampling error.
(c) Sample design must be viable in the context of funds available for the research study.
(d) Sample design must be such so that systematic bias can be controlled in a better way.
(e) Sample should be such that the results of the sample study can be applied, in general, for the universe with a reasonable level of confidence.

DIFFERENT TYPES OF SAMPLE DESIGNS

There are different types of sample designs based on two factors viz., the representation basis and the element selection technique. On the representation basis, the sample may be probability sampling or it may be non-probability sampling. Probability sampling is based on the concept of random selection, whereas non-probability sampling is 'non-random' sampling. On element selection basis, the sample may be either unrestricted or restricted. When each sample element is drawn individually from the population at large, then the sample so drawn is known as 'unrestricted sample', whereas all other forms of sampling are covered under the term 'restricted sampling'. The following chart exhibits the sample designs as explained above.

Thus, sample designs are basically of two types viz., non-probability sampling and probability sampling. We take up these two designs separately.

CHART SHOWING BASIC SAMPLING DESIGNS

Element selection technique ↓	Representation basis: Probability sampling	Representation basis: Non-probability sampling
Unrestricted sampling	Simple random sampling	Haphazard sampling or convenience sampling
Restricted sampling	Complex random sampling (such as cluster sampling, systematic sampling, stratified sampling etc.)	Purposive sampling (such as quota sampling, judgement sampling)

Fig. 4.1

Non-probability sampling: Non-probability sampling is that sampling procedure which does not afford any basis for estimating the probability that each item in the population has of being included in the sample. Non-probability sampling is also known by different names such as deliberate sampling, purposive sampling and judgement sampling. In this type of sampling, items for the sample are selected deliberately by the researcher; his choice concerning the items remains supreme. In other words, under non-probability sampling the organisers of the inquiry purposively choose the particular units of the universe for constituting a sample on the basis that the small mass that they so select out of a huge one will be typical or representative of the whole. For instance, if economic conditions of people living in a state are to be studied, a few towns and villages may be purposively selected for intensive study on the principle that they can be representative of the entire state. Thus, the judgement of the organisers of the study plays an important part in this sampling design.

In such a design, personal element has a great chance of entering into the selection of the sample. The investigator may select a sample which shall yield results favourable to his point of view and if that happens, the entire inquiry may get vitiated. Thus, there is always the danger of bias entering into this type of sampling technique. But in the investigators are impartial, work without bias and have the necessary experience so as to take sound judgement, the results obtained from an analysis of deliberately selected sample may be tolerably reliable. However, in such a sampling, there is no assurance that every element has some specifiable chance of being included. Sampling error in this type of sampling cannot be estimated and the element of bias, great or small, is always there. As such this sampling design in rarely adopted in large inquires of importance. However, in small inquiries and researches by individuals, this design may be adopted because of the relative advantage of time and money inherent in this method of sampling. *Quota sampling* is also an example of non-probability sampling. Under quota sampling the interviewers are simply given quotas to be filled from the different strata, with some restrictions on how they are to be filled. In other words, the actual selection of the items for the sample is left to the interviewer's discretion. This type of sampling is very convenient and is relatively inexpensive. But the samples so selected certainly do not possess the characteristic of random samples. Quota samples are essentially judgement samples and inferences drawn on their basis are not amenable to statistical treatment in a formal way.

Probability sampling: Probability sampling is also known as 'random sampling' or 'chance sampling'. Under this sampling design, every item of the universe has an equal chance of inclusion in the sample. It is, so to say, a lottery method in which individual units are picked up from the whole group not deliberately but by some mechanical process. Here it is blind chance alone that determines whether one item or the other is selected. The results obtained from probability or random sampling can be assured in terms of probability i.e., we can measure the errors of estimation or the significance of results obtained from a random sample, and this fact brings out the superiority of random sampling design over the deliberate sampling design. Random sampling ensures the law of Statistical Regularity which states that if on an average the sample chosen is a random one, the sample will have the same composition and characteristics as the universe. This is the reason why random sampling is considered as the best technique of selecting a representative sample.

Random sampling from a finite population refers to that method of sample selection which gives each possible sample combination an equal probability of being picked up and each item in the entire population to have an equal chance of being included in the sample. This applies to sampling without replacement i.e., once an item is selected for the sample, it cannot appear in the sample again (Sampling with replacement is used less frequently in which procedure the element selected for the sample is returned to the population before the next element is selected. In such a situation the same element could appear twice in the same sample before the second element is chosen). In brief, the implications of random sampling (or simple random sampling) are:

(a) It gives each element in the population an equal probability of getting into the sample; and all choices are independent of one another.

(b) It gives each possible sample combination an equal probability of being chosen.

Keeping this in view we can define a simple random sample (or simply a random sample) from a finite population as a sample which is chosen in such a way that each of the ${}^{N}C_{n}$ possible samples has the same probability, $1/{}^{N}C_{n}$, of being selected. To make it more clear we take a certain finite population consisting of six elements (say a, b, c, d, e, f) i.e., $N = 6$. Suppose that we want to take a sample of size $n = 3$ from it. Then there are ${}^{6}C_{3} = 20$ possible distinct samples of the required size, and they consist of the elements *abc*, *abd*, *abe*, *abf*, *acd*, *ace*, *acf*, *ade*, *adf*, *aef*, *bcd*, *bce*, *bcf*, *bde*, *bdf*, *bef*, *cde*, *cdf*, *cef*, and *def*. If we choose one of these samples in such a way that each has the probability 1/20 of being chosen, we will then call this a random sample.

HOW TO SELECT A RANDOM SAMPLE?

With regard to the question of how to take a random sample in actual practice, we could, in simple cases like the one above, write each of the possible samples on a slip of paper, mix these slips thoroughly in a container and then draw as a lottery either blindfolded or by rotating a drum or by any other similar device. Such a procedure is obviously impractical, if not altogether impossible in complex problems of sampling. In fact, the practical utility of such a method is very much limited.

Fortunately, we can take a random sample in a relatively easier way without taking the trouble of enlisting all possible samples on paper-slips as explained above. Instead of this, we can write the name of each element of a finite population on a slip of paper, put the slips of paper so prepared into a box or a bag and mix them thoroughly and then draw (without looking) the required number of slips for the sample one after the other without replacement. In doing so we must make sure that in

successive drawings each of the remaining elements of the population has the same chance of being selected. This procedure will also result in the same probability for each possible sample. We can verify this by taking the above example. Since we have a finite population of 6 elements and we want to select a sample of size 3, the probability of drawing any one element for our sample in the first draw is 3/6, the probability of drawing one more element in the second draw is 2/5, (the first element drawn is not replaced) and similarly the probability of drawing one more element in the third draw is 1/4. Since these draws are independent, the joint probability of the three elements which constitute our sample is the product of their individual probabilities and this works out to $3/6 \times 2/5 \times 1/4 = 1/20$. This verifies our earlier calculation.

Even this relatively easy method of obtaining a random sample can be simplified in actual practice by the use of random number tables. Various statisticians like Tippett, Yates, Fisher have prepared tables of random numbers which can be used for selecting a random sample. Generally, Tippett's random number tables are used for the purpose. Tippett gave10400 four figure numbers. He selected 41600 digits from the census reports and combined them into fours to give his random numbers which may be used to obtain a random sample.

We can illustrate the procedure by an example. First of all we reproduce the first thirty sets of Tippett's numbers

2952	6641	3992	9792	7979	5911
3170	5624	4167	9525	1545	1396
7203	5356	1300	2693	2370	7483
3408	2769	3563	6107	6913	7691
0560	5246	1112	9025	6008	8126

Suppose we are interested in taking a sample of 10 units from a population of 5000 units, bearing numbers from 3001 to 8000. We shall select 10 such figures from the above random numbers which are not less than 3001 and not greater than 8000. If we randomly decide to read the table numbers from left to right, starting from the first row itself, we obtain the following numbers: 6641, 3992, 7979, 5911, 3170, 5624, 4167, 7203, 5356, and 7483.

The units bearing the above serial numbers would then constitute our required random sample.

One may note that it is easy to draw random samples from finite populations with the aid of random number tables only when lists are available and items are readily numbered. But in some situations it is often impossible to proceed in the way we have narrated above. For example, if we want to estimate the mean height of trees in a forest, it would not be possible to number the trees, and choose random numbers to select a random sample. In such situations what we should do is to select some trees for the sample haphazardly without aim or purpose, and should treat the sample as a random sample for study purposes.

RANDOM SAMPLE FROM AN INFINITE UNIVERSE

So far we have talked about random sampling, keeping in view only the finite populations. But what about random sampling in context of infinite populations? It is relatively difficult to explain the concept of random sample from an infinite population. However, a few examples will show the basic characteristic of such a sample. Suppose we consider the 20 throws of a fair dice as a sample from the hypothetically infinite population which consists of the results of all possible throws of the dice. If

the probability of getting a particular number, say 1, is the same for each throw and the 20 throws are all independent, then we say that the sample is random. Similarly, it would be said to be sampling from an infinite population if we sample with replacement from a finite population and our sample would be considered as a random sample if in each draw all elements of the population have the same probability of being selected and successive draws happen to be independent. In brief, one can say that the selection of each item in a random sample from an infinite population is controlled by the same probabilities and that successive selections are independent of one another.

COMPLEX RANDOM SAMPLING DESIGNS

Probability sampling under restricted sampling techniques, as stated above, may result in complex random sampling designs. Such designs may as well be called 'mixed sampling designs' for many of such designs may represent a combination of probability and non-probability sampling procedures in selecting a sample. Some of the popular complex random sampling designs are as follows:

(i) Systematic sampling: In some instances, the most practical way of sampling is to select every *i*th item on a list. Sampling of this type is known as systematic sampling. An element of randomness is introduced into this kind of sampling by using random numbers to pick up the unit with which to start. For instance, if a 4 per cent sample is desired, the first item would be selected randomly from the first twenty-five and thereafter every 25th item would automatically be included in the sample. Thus, in systematic sampling only the first unit is selected randomly and the remaining units of the sample are selected at fixed intervals. Although a systematic sample is not a random sample in the strict sense of the term, but it is often considered reasonable to treat systematic sample as if it were a random sample.

Systematic sampling has certain plus points. It can be taken as an improvement over a simple random sample in as much as the systematic sample is spread more evenly over the entire population. It is an easier and less costlier method of sampling and can be conveniently used even in case of large populations. But there are certain dangers too in using this type of sampling. If there is a hidden periodicity in the population, systematic sampling will prove to be an inefficient method of sampling. For instance, every 25th item produced by a certain production process is defective. If we are to select a 4% sample of the items of this process in a systematic manner, we would either get all defective items or all good items in our sample depending upon the random starting position. If all elements of the universe are ordered in a manner representative of the total population, i.e., the population list is in random order, systematic sampling is considered equivalent to random sampling. But if this is not so, then the results of such sampling may, at times, not be very reliable. In practice, systematic sampling is used when lists of population are available and they are of considerable length.

(ii) Stratified sampling: If a population from which a sample is to be drawn does not constitute a homogeneous group, stratified sampling technique is generally applied in order to obtain a representative sample. Under stratified sampling the population is divided into several sub-populations that are individually more homogeneous than the total population (the different sub-populations are called 'strata') and then we select items from each stratum to constitute a sample. Since each stratum is more homogeneous than the total population, we are able to get more precise estimates for each stratum and by estimating more accurately each of the component parts, we get a better estimate of the whole. In brief, stratified sampling results in more reliable and detailed information.

The following three questions are highly relevant in the context of stratified sampling:

(a) How to form strata?

(b) How should items be selected from each stratum?

(c) How many items be selected from each stratum or how to allocate the sample size of each stratum?

Regarding the first question, we can say that the strata be formed on the basis of common characteristic(s) of the items to be put in each stratum. This means that various strata be formed in such a way as to ensure elements being most homogeneous within each stratum and most heterogeneous between the different strata. Thus, strata are purposively formed and are usually based on past experience and personal judgement of the researcher. One should always remember that careful consideration of the relationship between the characteristics of the population and the characteristics to be estimated are normally used to define the strata. At times, pilot study may be conducted for determining a more appropriate and efficient stratification plan. We can do so by taking small samples of equal size from each of the proposed strata and then examining the variances within and among the possible stratifications, we can decide an appropriate stratification plan for our inquiry.

In respect of the second question, we can say that the usual method, for selection of items for the sample from each stratum, resorted to is that of simple random sampling. Systematic sampling can be used if it is considered more appropriate in certain situations.

Regarding the third question, we usually follow the method of proportional allocation under which the sizes of the samples from the different strata are kept proportional to the sizes of the strata. That is, if P_i represents the proportion of population included in stratum i, and n represents the total sample size, the number of elements selected from stratum i is $n \,.\, P_i$. To illustrate it, let us suppose that we want a sample of size $n = 30$ to be drawn from a population of size $N = 8000$ which is divided into three strata of size $N_1 = 4000$, $N_2 = 2400$ and $N_3 = 1600$. Adopting proportional allocation, we shall get the sample sizes as under for the different strata:

For strata with $N_1 = 4000$, we have $P_1 = 4000/8000$

and hence $n_1 = n \,.\, P_1 = 30\ (4000/8000) = 15$

Similarly, for strata with $N_2 = 2400$, we have

$n_2 = n \,.\, P_2 = 30\ (2400/8000) = 9$, and

for strata with $N_3 = 1600$, we have

$n_3 = n \,.\, P_3 = 30\ (1600/8000) = 6$.

Thus, using proportional allocation, the sample sizes for different strata are 15, 9 and 6 respectively which is in proportion to the sizes of the strata viz., 4000 : 2400 : 1600. Proportional allocation is considered most efficient and an optimal design when the cost of selecting an item is equal for each stratum, there is no difference in within-stratum variances, and the purpose of sampling happens to be to estimate the population value of some characteristic. But in case the purpose happens to be to compare the differences among the strata, then equal sample selection from each stratum would be more efficient even if the strata differ in sizes. In cases where strata differ not only in size but also in variability and it is considered reasonable to take larger samples from the more variable strata and smaller samples from the less variable strata, we can then account for both (differences in stratum size and differences in stratum variability) by using disproportionate sampling design by requiring

$$n_1/N_1\sigma_1 = n_2/N_2\sigma_2 = \ldots\ldots\ldots = n_k/N_k\sigma_k$$

where $\sigma_1, \sigma_2, \ldots$ and σ_k denote the standard deviations of the k strata, $N_1, N_2, \ldots, N_k$ denote the sizes of the k strata and $n_1, n_2, \ldots, n_k$ denote the sample sizes of k strata. This is called '*optimum allocation*' in the context of disproportionate sampling. The allocation in such a situation results in the following formula for determining the sample sizes different strata:

$$n_i = \frac{n \cdot N_i \sigma_i}{N_1\sigma_1 + N_2\sigma_2 + \ldots + N_k\sigma_k} \qquad \text{for } i = 1, 2, \ldots \text{ and } k.$$

We may illustrate the use of this by an example.

Illustration 1

A population is divided into three strata so that $N_1 = 5000$, $N_2 = 2000$ and $N_3 = 3000$. Respective standard deviations are:

$$\sigma_1 = 15, \sigma_2 = 18 \text{ and } \sigma_3 = 5.$$

How should a sample of size $n = 84$ be allocated to the three strata, if we want optimum allocation using disproportionate sampling design?

Solution: Using the disproportionate sampling design for optimum allocation, the sample sizes for different strata will be determined as under:

Sample size for strata with $N_1 = 5000$

$$n_1 = \frac{84(5000)(15)}{(5000)(15) + (2000)(18) + (3000)(5)}$$

$$= 6300000/126000 = 50$$

Sample size for strata with $N_2 = 2000$

$$n_2 = \frac{84(2000)(18)}{(5000)(15) + (2000)(18) + (3000)(5)}$$

$$= 3024000/126000 = 24$$

Sample size for strata with $N_3 = 3000$

$$n_3 = \frac{84(3000)(5)}{(5000)(15) + (2000)(18) + (3000)(5)}$$

$$= 1260000/126000 = 10$$

In addition to differences in stratum size and differences in stratum variability, we may have differences in stratum sampling cost, then we can have cost optimal disproportionate sampling design by requiring

$$\frac{n_1}{N_1\sigma_1\sqrt{C_1}} = \frac{n_2}{N_2\sigma_2\sqrt{C_2}} = \ldots = \frac{n_k}{N_k\sigma_k\sqrt{C_k}}$$

where

C_1 = Cost of sampling in stratum 1

C_2 = Cost of sampling in stratum 2

C_k = Cost of sampling in stratum k

and all other terms remain the same as explained earlier. The allocation in such a situation results in the following formula for determining the sample sizes for different strata:

$$n_i = \frac{n \cdot N_i\sigma_i/\sqrt{C_i}}{N_1\sigma_1/\sqrt{C_1} + N_2\sigma_2/\sqrt{C_2} + \ldots + N_k\sigma_k/\sqrt{C_k}} \text{ for } i = 1, 2, \ldots, k$$

It is not necessary that stratification be done keeping in view a single characteristic. Populations are often stratified according to several characteristics. For example, a system-wide survey designed to determine the attitude of students toward a new teaching plan, a state college system with 20 colleges might stratify the students with respect to class, sec and college. Stratification of this type is known as *cross-stratification*, and up to a point such stratification increases the reliability of estimates and is much used in opinion surveys.

From what has been stated above in respect of stratified sampling, we can say that the sample so constituted is the result of successive application of purposive (involved in stratification of items) and random sampling methods. As such it is an example of mixed sampling. The procedure wherein we first have stratification and then simple random sampling is known as stratified random sampling.

(iii) Cluster sampling: If the total area of interest happens to be a big one, a convenient way in which a sample can be taken is to divide the area into a number of smaller non-overlapping areas and then to randomly select a number of these smaller areas (usually called clusters), with the ultimate sample consisting of all (or samples of) units in these small areas or clusters.

Thus in cluster sampling the total population is divided into a number of relatively small subdivisions which are themselves clusters of still smaller units and then some of these clusters are randomly selected for inclusion in the overall sample. Suppose we want to estimate the proportion of machine-parts in an inventory which are defective. Also assume that there are 20000 machine parts in the inventory at a given point of time, stored in 400 cases of 50 each. Now using a cluster sampling, we would consider the 400 cases as clusters and randomly select 'n' cases and examine all the machine-parts in each randomly selected case.

Cluster sampling, no doubt, reduces cost by concentrating surveys in selected clusters. But certainly it is less precise than random sampling. There is also not as much information in 'n' observations within a cluster as there happens to be in 'n' randomly drawn observations. Cluster sampling is used only because of the economic advantage it possesses; estimates based on cluster samples are usually more reliable per unit cost.

(iv) Area sampling: If clusters happen to be some geographic subdivisions, in that case cluster sampling is better known as area sampling. In other words, cluster designs, where the primary sampling unit represents a cluster of units based on geographic area, are distinguished as area sampling. The plus and minus points of cluster sampling are also applicable to area sampling.

(v) Multi-stage sampling: Multi-stage sampling is a further development of the principle of cluster sampling. Suppose we want to investigate the working efficiency of nationalised banks in India and we want to take a sample of few banks for this purpose. The first stage is to select large primary

sampling unit such as states in a country. Then we may select certain districts and interview all banks in the chosen districts. This would represent a two-stage sampling design with the ultimate sampling units being clusters of districts.

If instead of taking a census of all banks within the selected districts, we select certain towns and interview all banks in the chosen towns. This would represent a three-stage sampling design. If instead of taking a census of all banks within the selected towns, we randomly sample banks from each selected town, then it is a case of using a four-stage sampling plan. If we select randomly at all stages, we will have what is known as 'multi-stage random sampling design'.

Ordinarily multi-stage sampling is applied in big inquires extending to a considerable large geographical area, say, the entire country. There are two advantages of this sampling design viz., (a) It is easier to administer than most single stage designs mainly because of the fact that sampling frame under multi-stage sampling is developed in partial units. (b) A large number of units can be sampled for a given cost under multistage sampling because of sequential clustering, whereas this is not possible in most of the simple designs.

(vi) Sampling with probability proportional to size: In case the cluster sampling units do not have the same number or approximately the same number of elements, it is considered appropriate to use a random selection process where the probability of each cluster being included in the sample is proportional to the size of the cluster. For this purpose, we have to list the number of elements in each cluster irrespective of the method of ordering the cluster. Then we must sample systematically the appropriate number of elements from the cumulative totals. The actual numbers selected in this way do not refer to individual elements, but indicate which clusters and how many from the cluster are to be selected by simple random sampling or by systematic sampling. The results of this type of sampling are equivalent to those of a simple random sample and the method is less cumbersome and is also relatively less expensive. We can illustrate this with the help of an example.

Illustration 2

The following are the number of departmental stores in 15 cities: 35, 17, 10, 32, 70, 28, 26, 19, 26, 66, 37, 44, 33, 29 and 28. If we want to select a sample of 10 stores, using cities as clusters and selecting within clusters proportional to size, how many stores from each city should be chosen? (Use a starting point of 10).

Solution: Let us put the information as under (Table 4.1):

Since in the given problem, we have 500 departmental stores from which we have to select a sample of 10 stores, the appropriate sampling interval is 50. As we have to use the starting point of 10*, so we add successively increments of 50 till 10 numbers have been selected. The numbers, thus, obtained are: 10, 60, 110, 160, 210, 260, 310, 360, 410 and 460 which have been shown in the last column of the table (Table 4.1) against the concerning cumulative totals. From this we can say that two stores should be selected randomly from city number five and one each from city number 1, 3, 7, 9, 10, 11, 12, and 14. This sample of 10 stores is the sample with probability proportional to size.

*If the starting point is not mentioned, then the same can randomly be selected.

Table 4.1

City number	No. of departmental stores	Cumulative total	Sample	
1	35	35	10	
2	17	52		
3	10	62	60	
4	32	94		
5	70	164	110	160
6	28	192		
7	26	218	210	
8	19	237		
9	26	263	260	
10	66	329	310	
11	37	366	360	
12	44	410	410	
13	33	443		
14	29	472	460	
15	28	500		

(vii) **Sequential sampling:** This sampling design is some what complex sample design. The ultimate size of the sample under this technique is not fixed in advance, but is determined according to mathematical decision rules on the basis of information yielded as survey progresses. This is usually adopted in case of acceptance sampling plan in context of statistical quality control. When a particular lot is to be accepted or rejected on the basis of a single sample, it is known as single sampling; when the decision is to be taken on the basis of two samples, it is known as double sampling and in case the decision rests on the basis of more than two samples but the number of samples is certain and decided in advance, the sampling is known as multiple sampling. But when the number of samples is more than two but it is neither certain nor decided in advance, this type of system is often referred to as sequential sampling. Thus, in brief, we can say that in sequential sampling, one can go on taking samples one after another as long as one desires to do so.

CONCLUSION

From a brief description of the various sample designs presented above, we can say that normally one should resort to simple random sampling because under it bias is generally eliminated and the sampling error can be estimated. But purposive sampling is considered more appropriate when the universe happens to be small and a known characteristic of it is to be studied intensively. There are situations in real life under which sample designs other than simple random samples may be considered better (say easier to obtain, cheaper or more informative) and as such the same may be used. In a situation when random sampling is not possible, then we have to use necessarily a sampling design other than random sampling. At times, several methods of sampling may well be used in the same study.

Questions

1. What do you mean by 'Sample Design'? What points should be taken into consideration by a researcher in developing a sample design for this research project.
2. How would you differentiate between simple random sampling and complex random sampling designs? Explain clearly giving examples.
3. Why probability sampling is generally preferred in comparison to non-probability sampling? Explain the procedure of selecting a simple random sample.
4. Under what circumstances stratified random sampling design is considered appropriate? How would you select such sample? Explain by means of an example.
5. Distinguish between:
   (a) Restricted and unrestricted sampling;
   (b) Convenience and purposive sampling;
   (c) Systematic and stratified sampling;
   (d) Cluster and area sampling.
6. Under what circumstances would you recommend:
   (a) A probability sample?
   (b) A non-probability sample?
   (c) A stratified sample?
   (d) A cluster sample?
7. Explain and illustrate the procedure of selecting a random sample.
8. "A systematic bias results from errors in the sampling procedures". What do you mean by such a systematic bias? Describe the important causes responsible for such a bias.
9. (a) The following are the number of departmental stores in 10 cities: 35, 27, 24, 32, 42, 30, 34, 40, 29 and 38. If we want to select a sample of 15 stores using cities as clusters and selecting within clusters proportional to size, how many stores from each city should be chosen? (Use a starting point of 4).
   (b) What sampling design might be used to estimate the weight of a group of men and women?
10. A certain population is divided into five strata so that $N_1 = 2000, N_2 = 2000, N_3 = 1800, N_4 = 1700$, and $N_5 = 2500$. Respective standard deviations are: $\sigma_1 = 1.6, \sigma_2 = 2.0, \sigma_3 = 4.4, \sigma_4 = 4.8, \sigma_5 = 6.0$ and further the expected sampling cost in the first two strata is Rs 4 per interview and in the remaining three strata the sampling cost is Rs 6 per interview. How should a sample of size $n = 226$ be allocated to five strata if we adopt proportionate sampling design; if we adopt disproportionate sampling design considering (i) only the differences in stratum variability (ii) differences in stratum variability as well as the differences in stratum sampling costs.

5

Measurement and Scaling Techniques

MEASUREMENT IN RESEARCH

In our daily life we are said to measure when we use some yardstick to determine weight, height, or some other feature of a physical object. We also measure when we judge how well we like a song, a painting or the personalities of our friends. We, thus, measure physical objects as well as abstract concepts. Measurement is a relatively complex and demanding task, specially so when it concerns qualitative or abstract phenomena. By measurement we mean the process of assigning numbers to objects or observations, the level of measurement being a function of the rules under which the numbers are assigned.

It is easy to assign numbers in respect of properties of some objects, but it is relatively difficult in respect of others. For instance, measuring such things as social conformity, intelligence, or marital adjustment is much less obvious and requires much closer attention than measuring physical weight, biological age or a person's financial assets. In other words, properties like weight, height, etc., can be measured directly with some standard unit of measurement, but it is not that easy to measure properties like motivation to succeed, ability to stand stress and the like. We can expect high accuracy in measuring the length of pipe with a yard stick, but if the concept is abstract and the measurement tools are not standardized, we are less confident about the accuracy of the results of measurement.

Technically speaking, measurement is a process of mapping aspects of a domain onto other aspects of a range according to some rule of correspondence. In measuring, we devise some form of scale in the range (in terms of set theory, range may refer to some set) and then transform or map the properties of objects from the domain (in terms of set theory, domain may refer to some other set) onto this scale. For example, in case we are to find the male to female attendance ratio while conducting a study of persons who attend some show, then we may tabulate those who come to the show according to sex. In terms of set theory, this process is one of mapping the observed physical properties of those coming to the show (the domain) on to a sex classification (the range). The rule of correspondence is: If the object in the domain appears to be male, assign to "0" and if female assign to "1". Similarly, we can record a person's marital status as 1, 2, 3 or 4, depending on whether

the person is single, married, widowed or divorced. We can as well record "Yes or No" answers to a question as "0" and "1" (or as 1 and 2 or perhaps as 59 and 60). In this artificial or nominal way, categorical data (qualitative or descriptive) can be made into numerical data and if we thus code the various categories, we refer to the numbers we record as nominal data. *Nominal data* are numerical in name only, because they do not share any of the properties of the numbers we deal in ordinary arithmetic. For instance if we record marital status as 1, 2, 3, or 4 as stated above, we cannot write $4 > 2$ or $3 < 4$ and we cannot write $3 - 1 = 4 - 2$, $1 + 3 = 4$ or $4 \div 2 = 2$.

In those situations when we cannot do anything except set up inequalities, we refer to the data as *ordinal data*. For instance, if one mineral can scratch another, it receives a higher hardness number and on Mohs' scale the numbers from 1 to 10 are assigned respectively to talc, gypsum, calcite, fluorite, apatite, feldspar, quartz, topaz, sapphire and diamond. With these numbers we can write $5 > 2$ or $6 < 9$ as apatite is harder than gypsum and feldspar is softer than sapphire, but we cannot write for example $10 - 9 = 5 - 4$, because the difference in hardness between diamond and sapphire is actually much greater than that between apatite and fluorite. It would also be meaningless to say that topaz is twice as hard as fluorite simply because their respective hardness numbers on Mohs' scale are 8 and 4. The greater than symbol (i.e., >) in connection with ordinal data may be used to designate "happier than" "preferred to" and so on.

When in addition to setting up inequalities we can also form differences, we refer to the data as *interval data*. Suppose we are given the following temperature readings (in degrees Fahrenheit): 58°, 63°, 70°, 95°, 110°, 126° and 135°. In this case, we can write $100° > 70°$ or $95° < 135°$ which simply means that 110° is warmer than 70° and that 95° is cooler than 135°. We can also write for example $95° - 70° = 135° - 110°$, since equal temperature differences are equal in the sense that the same amount of heat is required to raise the temperature of an object from 70° to 95° or from 110° to 135°. On the other hand, it would not mean much if we said that 126° is twice as hot as 63°, even though $126° \div 63° = 2$. To show the reason, we have only to change to the centigrade scale, where the first temperature becomes $5/9\ (126 - 32) = 52°$, the second temperature becomes $5/9\ (63 - 32) = 17°$ and the first figure is now more than three times the second. This difficulty arises from the fact that Fahrenheit and Centigrade scales both have artificial origins (zeros) i.e., the number 0 of neither scale is indicative of the absence of whatever quantity we are trying to measure.

When in addition to setting up inequalities and forming differences we can also form quotients (i.e., when we can perform all the customary operations of mathematics), we refer to such data as *ratio data*. In this sense, ratio data includes all the usual measurement (or determinations) of length, height, money amounts, weight, volume, area, pressures etc.

The above stated distinction between nominal, ordinal, interval and ratio data is important for the nature of a set of data may suggest the use of particular statistical techniques*. A researcher has to be quite alert about this aspect while measuring properties of objects or of abstract concepts.

* When data can be measured in units which are interchangeable e.g., weights (by ratio scales), temperatures (by interval scales), that data is said to be parametric and can be subjected to most kinds of statistical and mathematical processes. But when data is measured in units which are not interchangeable, e.g., product preferences (by ordinal scales), the data is said to be non-parametric and is susceptible only to a limited extent to mathematical and statistical treatment.

MEASUREMENT SCALES

From what has been stated above, we can write that scales of measurement can be considered in terms of their mathematical properties. The most widely used classification of measurement scales are: (a) nominal scale; (b) ordinal scale; (c) interval scale; and (d) ratio scale.

(a) Nominal scale: Nominal scale is simply a system of assigning number symbols to events in order to label them. The usual example of this is the assignment of numbers of basketball players in order to identify them. Such numbers cannot be considered to be associated with an ordered scale for their order is of no consequence; the numbers are just convenient labels for the particular class of events and as such have no quantitative value. Nominal scales provide convenient ways of keeping track of people, objects and events. One cannot do much with the numbers involved. For example, one cannot usefully average the numbers on the back of a group of football players and come up with a meaningful value. Neither can one usefully compare the numbers assigned to one group with the numbers assigned to another. The counting of members in each group is the only possible arithmetic operation when a nominal scale is employed. Accordingly, we are restricted to use mode as the measure of central tendency. There is no generally used measure of dispersion for nominal scales. Chi-square test is the most common test of statistical significance that can be utilized, and for the measures of correlation, the contingency coefficient can be worked out.

Nominal scale is the least powerful level of measurement. It indicates no order or distance relationship and has no arithmetic origin. A nominal scale simply describes differences between things by assigning them to categories. Nominal data are, thus, counted data. The scale wastes any information that we may have about varying degrees of attitude, skills, understandings, etc. In spite of all this, nominal scales are still very useful and are widely used in surveys and other *ex-post-facto* research when data are being classified by major sub-groups of the population.

(b) Ordinal scale: The lowest level of the ordered scale that is commonly used is the ordinal scale. The ordinal scale places events in order, but there is no attempt to make the intervals of the scale equal in terms of some rule. Rank orders represent ordinal scales and are frequently used in research relating to qualitative phenomena. A student's rank in his graduation class involves the use of an ordinal scale. One has to be very careful in making statement about scores based on ordinal scales. For instance, if Ram's position in his class is 10 and Mohan's position is 40, it cannot be said that Ram's position is four times as good as that of Mohan. The statement would make no sense at all. Ordinal scales only permit the ranking of items from highest to lowest. Ordinal measures have no absolute values, and the real differences between adjacent ranks may not be equal. All that can be said is that one person is higher or lower on the scale than another, but more precise comparisons cannot be made.

Thus, the use of an ordinal scale implies a statement of 'greater than' or 'less than' (an equality statement is also acceptable) without our being able to state how much greater or less. The real difference between ranks 1 and 2 may be more or less than the difference between ranks 5 and 6. Since the numbers of this scale have only a rank meaning, the appropriate measure of central tendency is the median. A percentile or quartile measure is used for measuring dispersion. Correlations are restricted to various rank order methods. Measures of statistical significance are restricted to the non-parametric methods.

(c) Interval scale: In the case of interval scale, the intervals are adjusted in terms of some rule that has been established as a basis for making the units equal. The units are equal only in so far as one

accepts the assumptions on which the rule is based. Interval scales can have an arbitrary zero, but it is not possible to determine for them what may be called an absolute zero or the unique origin. The primary limitation of the interval scale is the lack of a true zero; it does not have the capacity to measure the complete absence of a trait or characteristic. The Fahrenheit scale is an example of an interval scale and shows similarities in what one can and cannot do with it. One can say that an increase in temperature from 30° to 40° involves the same increase in temperature as an increase from 60° to 70°, but one cannot say that the temperature of 60° is twice as warm as the temperature of 30° because both numbers are dependent on the fact that the zero on the scale is set arbitrarily at the temperature of the freezing point of water. The ratio of the two temperatures, 30° and 60°, means nothing because zero is an arbitrary point.

Interval scales provide more powerful measurement than ordinal scales for interval scale also incorporates the concept of equality of interval. As such more powerful statistical measures can be used with interval scales. Mean is the appropriate measure of central tendency, while standard deviation is the most widely used measure of dispersion. Product moment correlation techniques are appropriate and the generally used tests for statistical significance are the 't' test and 'F' test.

(d) Ratio scale: Ratio scales have an absolute or true zero of measurement. The term 'absolute zero' is not as precise as it was once believed to be. We can conceive of an absolute zero of length and similarly we can conceive of an absolute zero of time. For example, the zero point on a centimeter scale indicates the complete absence of length or height. But an absolute zero of temperature is theoretically unobtainable and it remains a concept existing only in the scientist's mind. The number of minor traffic-rule violations and the number of incorrect letters in a page of type script represent scores on ratio scales. Both these scales have absolute zeros and as such all minor traffic violations and all typing errors can be assumed to be equal in significance. With ratio scales involved one can make statements like "Jyoti's" typing performance was twice as good as that of "Reetu." The ratio involved does have significance and facilitates a kind of comparison which is not possible in case of an interval scale.

Ratio scale represents the actual amounts of variables. Measures of physical dimensions such as weight, height, distance, etc. are examples. Generally, all statistical techniques are usable with ratio scales and all manipulations that one can carry out with real numbers can also be carried out with ratio scale values. Multiplication and division can be used with this scale but not with other scales mentioned above. Geometric and harmonic means can be used as measures of central tendency and coefficients of variation may also be calculated.

Thus, proceeding from the nominal scale (the least precise type of scale) to ratio scale (the most precise), relevant information is obtained increasingly. If the nature of the variables permits, the researcher should use the scale that provides the most precise description. Researchers in physical sciences have the advantage to describe variables in ratio scale form but the behavioural sciences are generally limited to describe variables in interval scale form, a less precise type of measurement.

Sources of Error in Measurement

Measurement should be precise and unambiguous in an ideal research study. This objective, however, is often not met with in entirety. As such the researcher must be aware about the sources of error in measurement. The following are the possible sources of error in measurement.

(a) **Respondent:** At times the respondent may be reluctant to express strong negative feelings or it is just possible that he may have very little knowledge but may not admit his ignorance. All this reluctance is likely to result in an interview of 'guesses.' Transient factors like fatigue, boredom, anxiety, etc. may limit the ability of the respondent to respond accurately and fully.

(b) **Situation:** Situational factors may also come in the way of correct measurement. Any condition which places a strain on interview can have serious effects on the interviewer-respondent rapport. For instance, if someone else is present, he can distort responses by joining in or merely by being present. If the respondent feels that anonymity is not assured, he may be reluctant to express certain feelings.

(c) **Measurer:** The interviewer can distort responses by rewording or reordering questions. His behaviour, style and looks may encourage or discourage certain replies from respondents. Careless mechanical processing may distort the findings. Errors may also creep in because of incorrect coding, faulty tabulation and/or statistical calculations, particularly in the data-analysis stage.

(d) **Instrument:** Error may arise because of the defective measuring instrument. The use of complex words, beyond the comprehension of the respondent, ambiguous meanings, poor printing, inadequate space for replies, response choice omissions, etc. are a few things that make the measuring instrument defective and may result in measurement errors. Another type of instrument deficiency is the poor sampling of the universe of items of concern.

Researcher must know that correct measurement depends on successfully meeting all of the problems listed above. He must, to the extent possible, try to eliminate, neutralize or otherwise deal with all the possible sources of error so that the final results may not be contaminated.

Tests of Sound Measurement

Sound measurement must meet the tests of validity, reliability and practicality. In fact, these are the three major considerations one should use in evaluating a measurement tool. "Validity refers to the extent to which a test measures what we actually wish to measure. Reliability has to do with the accuracy and precision of a measurement procedure ... Practicality is concerned with a wide range of factors of economy, convenience, and interpretability ..."[1] We briefly take up the relevant details concerning these tests of sound measurement.

1. Test of Validity[*]

Validity is the most critical criterion and indicates the degree to which an instrument measures what it is supposed to measure. Validity can also be thought of as utility. In other words, validity is the extent to which differences found with a measuring instrument reflect true differences among those being tested. But the question arises: how can one determine validity without direct confirming knowledge? The answer may be that we seek other relevant evidence that confirms the answers we have found with our measuring tool. What is relevant, evidence often depends upon the nature of the

[1] Robert L. Thorndike and Elizabeth Hagen: *Measurement and Evaluation in Psychology and Education*, 3rd Ed., p. 162.

[*] Two forms of validity are usually mentioned in research literature viz., the external validity and the internal validity. External validity of research findings is their generalizability to populations, settings, treatment variables and measurement variables. We shall talk about it in the context of significance tests later on. The internal validity of a research design is its ability to measure what it aims to measure. We shall deal with this validity only in the present chapter.

research problem and the judgement of the researcher. But one can certainly consider three types of validity in this connection: (i) Content validity; (ii) Criterion-related validity and (iii) Construct validity.

(i) *Content validity* is the extent to which a measuring instrument provides adequate coverage of the topic under study. If the instrument contains a representative sample of the universe, the content validity is good. Its determination is primarily judgemental and intuitive. It can also be determined by using a panel of persons who shall judge how well the measuring instrument meets the standards, but there is no numerical way to express it.

(ii) *Criterion-related validity* relates to our ability to predict some outcome or estimate the existence of some current condition. This form of validity reflects the success of measures used for some empirical estimating purpose. The concerned criterion must possess the following qualities:

Relevance: (A criterion is relevant if it is defined in terms we judge to be the proper measure.)

Freedom from bias: (Freedom from bias is attained when the criterion gives each subject an equal opportunity to score well.)

Reliability: (A reliable criterion is stable or reproducible.)

Availability: (The information specified by the criterion must be available.)

In fact, a Criterion-related validity is a broad term that actually refers to (i) *Predictive validity* and *(ii) Concurrent validity*. The former refers to the usefulness of a test in predicting some future performance whereas the latter refers to the usefulness of a test in closely relating to other measures of known validity. Criterion-related validity is expressed as the coefficient of correlation between test scores and some measure of future performance or between test scores and scores on another measure of known validity.

(iii) *Construct validity* is the most complex and abstract. A measure is said to possess construct validity to the degree that it confirms to predicted correlations with other theoretical propositions. Construct validity is the degree to which scores on a test can be accounted for by the explanatory constructs of a sound theory. For determining construct validity, we associate a set of other propositions with the results received from using our measurement instrument. If measurements on our devised scale correlate in a predicted way with these other propositions, we can conclude that there is some construct validity.

If the above stated criteria and tests are met with, we may state that our measuring instrument is valid and will result in correct measurement; otherwise we shall have to look for more information and/or resort to exercise of judgement.

2. Test of Reliability

The test of reliability is another important test of sound measurement. A measuring instrument is reliable if it provides consistent results. Reliable measuring instrument does contribute to validity, but a reliable instrument need not be a valid instrument. For instance, a scale that consistently overweighs objects by five kgs., is a reliable scale, but it does not give a valid measure of weight. But the other way is not true i.e., a valid instrument is always reliable. Accordingly reliability is not as valuable as validity, but it is easier to assess reliability in comparison to validity. If the quality of reliability is satisfied by an instrument, then while using it we can be confident that the transient and situational factors are not interfering.

Two aspects of reliability viz., stability and equivalence deserve special mention. The *stability aspect* is concerned with securing consistent results with repeated measurements of the same person and with the same instrument. We usually determine the degree of stability by comparing the results of repeated measurements. The *equivalence aspect* considers how much error may get introduced by different investigators or different samples of the items being studied. A good way to test for the equivalence of measurements by two investigators is to compare their observations of the same events. Reliability can be improved in the following two ways:

(i) By standardising the conditions under which the measurement takes place i.e., we must ensure that external sources of variation such as boredom, fatigue, etc., are minimised to the extent possible. That will improve stability aspect.

(ii) By carefully designed directions for measurement with no variation from group to group, by using trained and motivated persons to conduct the research and also by broadening the sample of items used. This will improve equivalence aspect.

3. Test of Practicality

The practicality characteristic of a measuring instrument can be judged in terms of economy, convenience and interpretability. From the operational point of view, the measuring instrument ought to be practical i.e., it should be economical, convenient and interpretable. *Economy* consideration suggests that some trade-off is needed between the ideal research project and that which the budget can afford. The length of measuring instrument is an important area where economic pressures are quickly felt. Although more items give greater reliability as stated earlier, but in the interest of limiting the interview or observation time, we have to take only few items for our study purpose. Similarly, data-collection methods to be used are also dependent at times upon economic factors. *Convenience* test suggests that the measuring instrument should be easy to administer. For this purpose one should give due attention to the proper layout of the measuring instrument. For instance, a questionnaire, with clear instructions (illustrated by examples), is certainly more effective and easier to complete than one which lacks these features. *Interpretability* consideration is specially important when persons other than the designers of the test are to interpret the results. The measuring instrument, in order to be interpretable, must be supplemented by (a) detailed instructions for administering the test; (b) scoring keys; (c) evidence about the reliability and (d) guides for using the test and for interpreting results.

TECHNIQUE OF DEVELOPING MEASUREMENT TOOLS

The technique of developing measurement tools involves a four-stage process, consisting of the following:

(a) Concept development;
(b) Specification of concept dimensions;
(c) Selection of indicators; and
(d) Formation of index.

The first and foremost step is that of *concept development* which means that the researcher should arrive at an understanding of the major concepts pertaining to his study. This step of concept

development is more apparent in theoretical studies than in the more pragmatic research, where the fundamental concepts are often already established.

The second step requires the researcher to specify the *dimensions of the concepts* that he developed in the first stage. This task may either be accomplished by deduction i.e., by adopting a more or less intuitive approach or by empirical correlation of the individual dimensions with the total concept and/or the other concepts. For instance, one may think of several dimensions such as product reputation, customer treatment, corporate leadership, concern for individuals, sense of social responsibility and so forth when one is thinking about the image of a certain company.

Once the dimensions of a concept have been specified, the researcher must *develop indicators* for measuring each concept element. Indicators are specific questions, scales, or other devices by which respondent's knowledge, opinion, expectation, etc., are measured. As there is seldom a perfect measure of a concept, the researcher should consider several alternatives for the purpose. The use of more than one indicator gives stability to the scores and it also improves their validity.

The last step is that of combining the various indicators into an index, i.e., *formation of an index*. When we have several dimensions of a concept or different measurements of a dimension, we may need to combine them into a single index. One simple way for getting an overall index is to provide scale values to the responses and then sum up the corresponding scores. Such an overall index would provide a better measurement tool than a single indicator because of the fact that an "individual indicator has only a probability relation to what we really want to know."[2] This way we must obtain an overall index for the various concepts concerning the research study.

Scaling

In research we quite often face measurement problem (since we want a valid measurement but may not obtain it), specially when the concepts to be measured are complex and abstract and we do not possess the standardised measurement tools. Alternatively, we can say that while measuring attitudes and opinions, we face the problem of their valid measurement. Similar problem may be faced by a researcher, of course in a lesser degree, while measuring physical or institutional concepts. As such we should study some procedures which may enable us to measure abstract concepts more accurately. This brings us to the study of scaling techniques.

Meaning of Scaling

Scaling describes the procedures of assigning numbers to various degrees of opinion, attitude and other concepts. This can be done in two ways viz., (i) making a judgement about some characteristic of an individual and then placing him directly on a scale that has been defined in terms of that characteristic and (ii) constructing questionnaires in such a way that the score of individual's responses assigns him a place on a scale. It may be stated here that a scale is a continuum, consisting of the highest point (in terms of some characteristic e.g., preference, favourableness, etc.) and the lowest point along with several intermediate points between these two extreme points. These scale-point positions are so related to each other that when the first point happens to be the highest point, the second point indicates a higher degree in terms of a given characteristic as compared to the third

[2] Lazersfeld, *Evidence and Inference*, p. 112.

point and the third point indicates a higher degree as compared to the fourth and so on. Numbers for measuring the distinctions of degree in the attitudes/opinions are, thus, assigned to individuals corresponding to their scale-positions. All this is better understood when we talk about scaling technique(s). Hence the term 'scaling' is applied to the procedures for attempting to determine quantitative measures of subjective abstract concepts. Scaling has been defined as a "procedure for the assignment of numbers (or other symbols) to a property of objects in order to impart some of the characteristics of numbers to the properties in question."[3]

Scale Classification Bases

The number assigning procedures or the scaling procedures may be broadly classified on one or more of the following bases: (a) subject orientation; (b) response form; (c) degree of subjectivity; (d) scale properties; (e) number of dimensions and (f) scale construction techniques. We take up each of these separately.

(a) Subject orientation: Under it a scale may be designed to measure characteristics of the respondent who completes it or to judge the stimulus object which is presented to the respondent. In respect of the former, we presume that the stimuli presented are sufficiently homogeneous so that the between-stimuli variation is small as compared to the variation among respondents. In the latter approach, we ask the respondent to judge some specific object in terms of one or more dimensions and we presume that the between-respondent variation will be small as compared to the variation among the different stimuli presented to respondents for judging.

(b) Response form: Under this we may classify the scales as categorical and comparative. Categorical scales are also known as rating scales. These scales are used when a respondent scores some object without direct reference to other objects. Under comparative scales, which are also known as ranking scales, the respondent is asked to compare two or more objects. In this sense the respondent may state that one object is superior to the other or that three models of pen rank in order 1, 2 and 3. The essence of ranking is, in fact, a relative comparison of a certain property of two or more objects.

(c) Degree of subjectivity: With this basis the scale data may be based on whether we measure subjective personal preferences or simply make non-preference judgements. In the former case, the respondent is asked to choose which person he favours or which solution he would like to see employed, whereas in the latter case he is simply asked to judge which person is more effective in some aspect or which solution will take fewer resources without reflecting any personal preference.

(d) Scale properties: Considering scale properties, one may classify the scales as nominal, ordinal, interval and ratio scales. Nominal scales merely classify without indicating order, distance or unique origin. Ordinal scales indicate magnitude relationships of 'more than' or 'less than', but indicate no distance or unique origin. Interval scales have both order and distance values, but no unique origin. Ratio scales possess all these features.

(e) Number of dimensions: In respect of this basis, scales can be classified as 'unidimensional' and 'multidimensional' scales. Under the former we measure only one attribute of the respondent or object, whereas multidimensional scaling recognizes that an object might be described better by using the concept of an attribute space of 'n' dimensions, rather than a single-dimension continuum.

[3] Bernard S. Phillips, *Social Research Strategy and Tactics*, 2nd ed., p. 205.

(f) Scale construction techniques: Following are the five main techniques by which scales can be developed.

(i) *Arbitrary approach:* It is an approach where scale is developed on *ad hoc* basis. This is the most widely used approach. It is presumed that such scales measure the concepts for which they have been designed, although there is little evidence to support such an assumption.

(ii) *Consensus approach:* Here a panel of judges evaluate the items chosen for inclusion in the instrument in terms of whether they are relevant to the topic area and unambiguous in implication.

(iii) *Item analysis approach:* Under it a number of individual items are developed into a test which is given to a group of respondents. After administering the test, the total scores are calculated for every one. Individual items are then analysed to determine which items discriminate between persons or objects with high total scores and those with low scores.

(iv) *Cumulative scales* are chosen on the basis of their conforming to some ranking of items with ascending and descending discriminating power. For instance, in such a scale the endorsement of an item representing an extreme position should also result in the endorsement of all items indicating a less extreme position.

(v) *Factor scales* may be constructed on the basis of intercorrelations of items which indicate that a common factor accounts for the relationship between items. This relationship is typically measured through factor analysis method.

Important Scaling Techniques

We now take up some of the important scaling techniques often used in the context of research specially in context of social or business research.

Rating scales: The rating scale involves qualitative description of a limited number of aspects of a thing or of traits of a person. When we use rating scales (or categorical scales), we judge an object in absolute terms against some specified criteria i.e., we judge properties of objects without reference to other similar objects. These ratings may be in such forms as "like-dislike", "above average, average, below average", or other classifications with more categories such as "like very much—like some what—neutral—dislike somewhat—dislike very much"; "excellent—good—average—below average—poor", "always—often—occasionally—rarely—never", and so on. There is no specific rule whether to use a two-points scale, three-points scale or scale with still more points. In practice, three to seven points scales are generally used for the simple reason that more points on a scale provide an opportunity for greater sensitivity of measurement.

Rating scale may be either a graphic rating scale or an itemized rating scale.

(i) *The graphic rating scale* is quite simple and is commonly used in practice. Under it the various points are usually put along the line to form a continuum and the rater indicates his rating by simply making a mark (such as ✓) at the appropriate point on a line that runs from one extreme to the other. Scale-points with brief descriptions may be indicated along the line, their function being to assist the rater in performing his job. The following is an example of five-points graphic rating scale when we wish to ascertain people's liking or disliking any product:

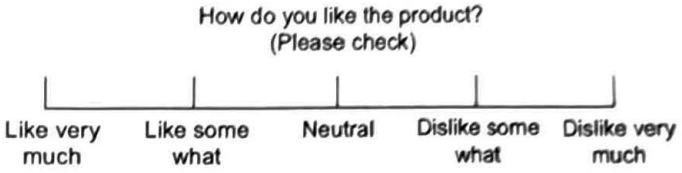


Fig. 5.1

This type of scale has several limitations. The respondents may check at almost any position along the line which fact may increase the difficulty of analysis. The meanings of the terms like "very much" and "some what" may depend upon respondent's frame of reference so much so that the statement might be challenged in terms of its equivalency. Several other rating scale variants (e.g., boxes replacing line) may also be used.

(ii) The *itemized rating scale* (also known as numerical scale) presents a series of statements from which a respondent selects one as best reflecting his evaluation. These statements are ordered progressively in terms of more or less of some property. An example of itemized scale can be given to illustrate it.

Suppose we wish to inquire as to how well does a worker get along with his fellow workers? In such a situation we may ask the respondent to select one, to express his opinion, from the following:

- He is almost always involved in some friction with a fellow worker.
- He is often at odds with one or more of his fellow workers.
- He sometimes gets involved in friction.
- He infrequently becomes involved in friction with others.
- He almost never gets involved in friction with fellow workers.

The chief merit of this type of scale is that it provides more information and meaning to the rater, and thereby increases reliability. This form is relatively difficult to develop and the statements may not say exactly what the respondent would like to express.

Rating scales have certain good points. The results obtained from their use compare favourably with alternative methods. They require less time, are interesting to use and have a wide range of applications. Besides, they may also be used with a large number of properties or variables. But their value for measurement purposes depends upon the assumption that the respondents can and do make good judgements. If the respondents are not very careful while rating, errors may occur. Three types of errors are common viz., the error of leniency, the error of central tendency and the error of hallo effect. The error of leniency occurs when certain respondents are either easy raters or hard raters. When raters are reluctant to give extreme judgements, the result is the error of central tendency. The error of hallo effect or the systematic bias occurs when the rater carries over a generalised impression of the subject from one rating to another. This sort of error takes place when we conclude for example, that a particular report is good because we like its form or that someone is intelligent because he agrees with us or has a pleasing personality. In other words, hallo effect is likely to appear when the rater is asked to rate many factors, on a number of which he has no evidence for judgement.

Ranking scales: Under ranking scales (or comparative scales) we make relative judgements against other similar objects. The respondents under this method directly compare two or more objects and make choices among them. There are two generally used approaches of ranking scales viz.

(a) Method of paired comparisons: Under it the respondent can express his attitude by making a choice between two objects, say between a new flavour of soft drink and an established brand of drink. But when there are more than two stimuli to judge, the number of judgements required in a paired comparison is given by the formula:

$$N = \frac{n(n-1)}{2}$$

where N = number of judgements

n = number of stimuli or objects to be judged.

For instance, if there are ten suggestions for bargaining proposals available to a workers union, there are 45 paired comparisons that can be made with them. When N happens to be a big figure, there is the risk of respondents giving ill considered answers or they may even refuse to answer. We can reduce the number of comparisons per respondent either by presenting to each one of them only a sample of stimuli or by choosing a few objects which cover the range of attractiveness at about equal intervals and then comparing all other stimuli to these few standard objects. Thus, paired-comparison data may be treated in several ways. If there is substantial consistency, we will find that if X is preferred to Y, and Y to Z, then X will consistently be preferred to Z. If this is true, we may take the total number of preferences among the comparisons as the score for that stimulus.

It should be remembered that paired comparison provides ordinal data, but the same may be converted into an interval scale by the method of the *Law of Comparative Judgement* developed by L.L. Thurstone. This technique involves the conversion of frequencies of preferences into a table of proportions which are then transformed into Z matrix by referring to the table of area under the normal curve. J.P. Guilford in his book "Psychometric Methods" has given a procedure which is relatively easier. The method is known as the *Composite Standard Method* and can be illustrated as under:

Suppose there are four proposals which some union bargaining committee is considering. The committee wants to know how the union membership ranks these proposals. For this purpose a sample of 100 members might express the views as shown in the following table:

Table 5.1: Response Patterns of 100 Members' Paired Comparisons of 4 Suggestions for Union Bargaining Proposal Priorities

	Suggestion			
	A	*B*	*C*	*D*
A	–	65*	32	20
B	40	–	38	42
C	45	50	–	70
D	80	20	98	–
TOTAL:	165	135	168	132

Contd.

*Read as 65 members preferred suggestion *B* to suggestion *A*.

Rank order	2	3	1	4
M_p	0.5375	0.4625	0.5450	0.4550
Z_j	0.09	(–).09	0.11	(–).11
R_j	0.20	0.02	0.22	0.00

Comparing the total number of preferences for each of the four proposals, we find that C is the most popular, followed by A, B and D respectively in popularity. The rank order shown in the above table explains all this.

By following the composite standard method, we can develop an interval scale from the paired-comparison ordinal data given in the above table for which purpose we have to adopt the following steps in order:

(i) Using the data in the above table, we work out the column mean with the help of the formula given below:

$$M_p = \frac{C + .5(N)}{nN} = \frac{165 + .5(100)}{4(100)} = .5375$$

where

M_p = the mean proportion of the columns

C = the total number of choices for a given suggestion

n = number of stimuli (proposals in the given problem)

N = number of items in the sample.

The column means have been shown in the M_p row in the above table.

(ii) The Z values for the M_p are secured from the table giving the area under the normal curve. When the M_p value is less than .5, the Z value is negative and for all M_p values higher than .5, the Z values are positive.* These Z values are shown in Z_j row in the above table.

(iii) As the Z_j values represent an interval scale, zero is an arbitrary value. Hence we can eliminate negative scale values by giving the value of zero to the lowest scale value (this being (–).11 in our example which we shall take equal to zero) and then adding the absolute value of this lowest scale value to all other scale items. This scale has been shown in R_j row in the above table.

Graphically we can show this interval scale that we have derived from the paired-comparison data using the composite standard method as follows:

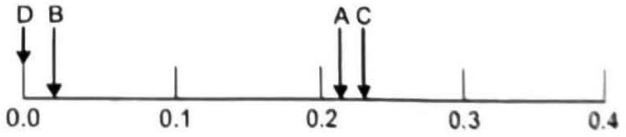


Fig. 5.2

*To use Normal curve area table for this sort of transformation, we must subtract 0.5 from all M_p values which exceed .5 to secure the values with which to enter the normal curve area table for which Z values can be obtained. For all M_p values of less than . 5 we must subtract all such values from 0.5 to secure the values with which to enter the normal curve area table for which Z values can be obtained but the Z values in this situation will be with negative sign.

(b) Method of rank order: Under this method of comparative scaling, the respondents are asked to rank their choices. This method is easier and faster than the method of paired comparisons stated above. For example, with 10 items it takes 45 pair comparisons to complete the task, whereas the method of rank order simply requires ranking of 10 items only. The problem of transitivity (such as *A* prefers to *B*, *B* to *C*, but *C* prefers to *A*) is also not there in case we adopt method of rank order. Moreover, a complete ranking at times is not needed in which case the respondents may be asked to rank only their first, say, four choices while the number of overall items involved may be more than four, say, it may be 15 or 20 or more. To secure a simple ranking of all items involved we simply total rank values received by each item. There are methods through which we can as well develop an interval scale of these data. But then there are limitations of this method. The first one is that data obtained through this method are ordinal data and hence rank ordering is an ordinal scale with all its limitations. Then there may be the problem of respondents becoming careless in assigning ranks particularly when there are many (usually more than 10) items.

Scale Construction Techniques

In social science studies, while measuring attitudes of the people we generally follow the technique of preparing the opinionnaire* (or attitude scale) in such a way that the score of the individual responses assigns him a place on a scale. Under this approach, the respondent expresses his agreement or disagreement with a number of statements relevant to the issue. While developing such statements, the researcher must note the following two points:

(i) That the statements must elicit responses which are psychologically related to the attitude being measured;

(ii) That the statements need be such that they discriminate not merely between extremes of attitude but also among individuals who differ slightly.

Researchers must as well be aware that inferring attitude from what has been recorded in opinionnaires has several limitations. People may conceal their attitudes and express socially acceptable opinions. They may not really know how they feel about a social issue. People may be unaware of their attitude about an abstract situation; until confronted with a real situation, they may be unable to predict their reaction. Even behaviour itself is at times not a true indication of attitude. For instance, when politicians kiss babies, their behaviour may not be a true expression of affection toward infants. Thus, there is no sure method of measuring attitude; we only try to measure the expressed opinion and then draw inferences from it about people's real feelings or attitudes.

With all these limitations in mind, psychologists and sociologists have developed several scale construction techniques for the purpose. The researcher should know these techniques so as to develop an appropriate scale for his own study. Some of the important approaches, along with the corresponding scales developed under each approach to measure attitude are as follows:

* An information form that attempts to measure the attitude or belief of an individual is known as opinionnaire.

Table 5.2: Different Scales for Measuring Attitudes of People

Name of the scale construction approach	*Name of the scale developed*
1. Arbitrary approach	Arbitrary scales
2. Consensus scale approach	Differential scales (such as Thurstone Differential scale)
3. Item analysis approach	Summated scales (such as Likert Scale)
4. Cumulative scale approach	Cumulative scales (such as Guttman's Scalogram)
5. Factor analysis approach	Factor scales (such as Osgood's Semantic Differential, Multi-dimensional Scaling, etc.)

A brief description of each of the above listed scales will be helpful.

Arbitrary Scales

Arbitrary scales are developed on *ad hoc* basis and are designed largely through the researcher's own subjective selection of items. The researcher first collects few statements or items which he believes are unambiguous and appropriate to a given topic. Some of these are selected for inclusion in the measuring instrument and then people are asked to check in a list the statements with which they agree.

The chief merit of such scales is that they can be developed very easily, quickly and with relatively less expense. They can also be designed to be highly specific and adequate. Because of these benefits, such scales are widely used in practice.

At the same time there are some limitations of these scales. The most important one is that we do not have objective evidence that such scales measure the concepts for which they have been developed. We have simply to rely on researcher's insight and competence.

Differential Scales (or Thurstone-type Scales)

The name of L.L. Thurstone is associated with differential scales which have been developed using consensus scale approach. Under such an approach the selection of items is made by a panel of judges who evaluate the items in terms of whether they are relevant to the topic area and unambiguous in implication. The detailed procedure is as under:

(a) The researcher gathers a large number of statements, usually twenty or more, that express various points of view toward a group, institution, idea, or practice (i.e., statements belonging to the topic area).

(b) These statements are then submitted to a panel of judges, each of whom arranges them in eleven groups or piles ranging from one extreme to another in position. Each of the judges is requested to place generally in the first pile the statements which he thinks are most unfavourable to the issue, in the second pile to place those statements which he thinks are next most unfavourable and he goes on doing so in this manner till in the eleventh pile he puts the statements which he considers to be the most favourable.

(c) This sorting by each judge yields a composite position for each of the items. In case of marked disagreement between the judges in assigning a position to an item, that item is discarded.

(d) For items that are retained, each is given its median scale value between one and eleven as established by the panel. In other words, the scale value of any one statement is computed as the 'median' position to which it is assigned by the group of judges.

(e) A final selection of statements is then made. For this purpose a sample of statements, whose median scores are spread evenly from one extreme to the other is taken. The statements so selected, constitute the final scale to be administered to respondents. The position of each statement on the scale is the same as determined by the judges.

After developing the scale as stated above, the respondents are asked during the administration of the scale to check the statements with which they agree. The median value of the statements that they check is worked out and this establishes their score or quantifies their opinion. It may be noted that in the actual instrument the statements are arranged in random order of scale value. If the values are valid and if the opinionnaire deals with only one attitude dimension, the typical respondent will choose one or several contiguous items (in terms of scale values) to reflect his views. However, at times divergence may occur when a statement appears to tap a different attitude dimension.

The Thurstone method has been widely used for developing differential scales which are utilised to measure attitudes towards varied issues like war, religion, etc. Such scales are considered most appropriate and reliable when used for measuring a single attitude. But an important deterrent to their use is the cost and effort required to develop them. Another weakness of such scales is that the values assigned to various statements by the judges may reflect their own attitudes. The method is not completely objective; it involves ultimately subjective decision process. Critics of this method also opine that some other scale designs give more information about the respondent's attitude in comparison to differential scales.

Summated Scales (or Likert-type Scales)

Summated scales (or Likert-type scales) are developed by utilizing the item analysis approach wherein a particular item is evaluated on the basis of how well it discriminates between those persons whose total score is high and those whose score is low. Those items or statements that best meet this sort of discrimination test are included in the final instrument.

Thus, summated scales consist of a number of statements which express either a favourable or unfavourable attitude towards the given object to which the respondent is asked to react. The respondent indicates his agreement or disagreement with each statement in the instrument. Each response is given a numerical score, indicating its favourableness or unfavourableness, and the scores are totalled to measure the respondent's attitude. In other words, the overall score represents the respondent's position on the continuum of favourable-unfavourableness towards an issue.

Most frequently used summated scales in the study of social attitudes follow the pattern devised by Likert. For this reason they are often referred to as Likert-type scales. In a Likert scale, the respondent is asked to respond to each of the statements in terms of several degrees, usually five degrees (but at times 3 or 7 may also be used) of agreement or disagreement. For example, when asked to express opinion whether one considers his job quite pleasant, the respondent may respond in any one of the following ways: (i) strongly agree, (ii) agree, (iii) undecided, (iv) disagree, (v) strongly disagree

We find that these five points constitute the scale. At one extreme of the scale there is strong agreement with the given statement and at the other, strong disagreement, and between them lie intermediate points. We may illustrate this as under:

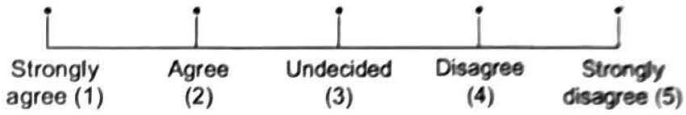


Fig. 5.3

Each point on the scale carries a score. Response indicating the least favourable degree of job satisfaction is given the least score (say 1) and the most favourable is given the highest score (say 5). These score—values are normally not printed on the instrument but are shown here just to indicate the scoring pattern. The Likert scaling technique, thus, assigns a scale value to each of the five responses. The same thing is done in respect of each and every statement in the instrument. This way the instrument yields a total score for each respondent, which would then measure the respondent's favourableness toward the given point of view. If the instrument consists of, say 30 statements, the following score values would be revealing.

$30 \times 5 = 150$ Most favourable response possible

$30 \times 3 = 90$ A neutral attitude

$30 \times 1 = 30$ Most unfavourable attitude.

The scores for any individual would fall between 30 and 150. If the score happens to be above 90, it shows favourable opinion to the given point of view, a score of below 90 would mean unfavourable opinion and a score of exactly 90 would be suggestive of a neutral attitude.

Procedure: The procedure for developing a Likert-type scale is as follows:

(i) As a first step, the researcher collects a large number of statements which are relevant to the attitude being studied and each of the statements expresses definite favourableness or unfavourableness to a particular point of view or the attitude and that the number of favourable and unfavourable statements is approximately equal.

(ii) After the statements have been gathered, a trial test should be administered to a number of subjects. In other words, a small group of people, from those who are going to be studied finally, are asked to indicate their response to each statement by checking one of the categories of agreement or disagreement using a five point scale as stated above.

(iii) The response to various statements are scored in such a way that a response indicative of the most favourable attitude is given the highest score of 5 and that with the most unfavourable attitude is given the lowest score, say, of 1.

(iv) Then the total score of each respondent is obtained by adding his scores that he received for separate statements.

(v) The next step is to array these total scores and find out those statements which have a high discriminatory power. For this purpose, the researcher may select some part of the highest and the lowest total scores, say the top 25 per cent and the bottom 25 per cent. These two extreme groups are interpreted to represent the most favourable and the least favourable attitudes and are used as criterion groups by which to evaluate individual statements. Thus

way we determine which statements consistently correlate with low favourability and which with high favourability.

(vi) Only those statements that correlate with the total test should be retained in the final instrument and all others must be discarded from it.

Advantages: The Likert-type scale has several advantages. Mention may be made of the important ones.

(a) It is relatively easy to construct the Likert-type scale in comparison to Thurstone-type scale because Likert-type scale can be performed without a panel of judges.

(b) Likert-type scale is considered more reliable because under it respondents answer each statement included in the instrument. As such it also provides more information and data than does the Thurstone-type scale.

(c) Each statement, included in the Likert-type scale, is given an empirical test for discriminating ability and as such, unlike Thurstone-type scale, the Likert-type scale permits the use of statements that are not manifestly related (to have a direct relationship) to the attitude being studied.

(d) Likert-type scale can easily be used in respondent-centred and stimulus-centred studies i.e., through it we can study how responses differ between people and how responses differ between stimuli.

(e) Likert-type scale takes much less time to construct, it is frequently used by the students of opinion research. Moreover, it has been reported in various research studies* that there is high degree of correlation between Likert-type scale and Thurstone-type scale.

Limitations: There are several limitations of the Likert-type scale as well. One important limitation is that, with this scale, we can simply examine whether respondents are more or less favourable to a topic, but we cannot tell how much more or less they are. There is no basis for belief that the five positions indicated on the scale are equally spaced. The interval between 'strongly agree' and 'agree', may not be equal to the interval between "agree" and "undecided". This means that Likert scale does not rise to a stature more than that of an ordinal scale, whereas the designers of Thurstone scale claim the Thurstone scale to be an interval scale. One further disadvantage is that often the total score of an individual respondent has little clear meaning since a given total score can be secured by a variety of answer patterns. It is unlikely that the respondent can validly react to a short statement on a printed form in the absence of real-life qualifying situations. Moreover, there "remains a possibility that people may answer according to what they think they should feel rather than how they do feel."[4] This particular weakness of the Likert-type scale is met by using a cumulative scale which we shall take up later in this chapter.

In spite of all the limitations, the Likert-type summated scales are regarded as the most useful in a situation wherein it is possible to compare the respondent's score with a distribution of scores from some well defined group. They are equally useful when we are concerned with a programme of

*A.L. Edwards and K.C. Kenney, "A comparison of the Thurstone and Likert techniques of attitude scale construction", *Journal of Applied Psychology*, 30, 72–83, 1946.

[4] John W. Best and James V. Kahn, "Research in Education", 5 ed., Prentice-Hall of India Pvt. Ltd., New Delhi, 1986, p. 183.

change or improvement in which case we can use the scales to measure attitudes before and after the programme of change or improvement in order to assess whether our efforts have had the desired effects. We can as well correlate scores on the scale to other measures without any concern for the absolute value of what is favourable and what is unfavourable. All this accounts for the popularity of Likert-type scales in social studies relating to measuring of attitudes.

Cumulative scales: Cumulative scales or Louis Guttman's scalogram analysis, like other scales, consist of series of statements to which a respondent expresses his agreement or disagreement. The special feature of this type of scale is that statements in it form a cumulative series. This, in other words, means that the statements are related to one another in such a way that an individual, who replies favourably to say item No. 3, also replies favourably to items No. 2 and 1, and one who replies favourably to item No. 4 also replies favourably to items No. 3, 2 and 1, and so on. This being so an individual whose attitude is at a certain point in a cumulative scale will answer favourably all the items on one side of this point, and answer unfavourably all the items on the other side of this point. The individual's score is worked out by counting the number of points concerning the number of statements he answers favourably. If one knows this total score, one can estimate as to how a respondent has answered individual statements constituting cumulative scales. The major scale of this type of cumulative scales is the Guttman's scalogram. We attempt a brief description of the same below.

The technique developed by Louis Guttman is known as scalogram analysis, or at times simply 'scale analysis'. Scalogram analysis refers to the procedure for determining whether a set of items forms a unidimensional scale. A scale is said to be unidimensional if the responses fall into a pattern in which endorsement of the item reflecting the extreme position results also in endorsing all items which are less extreme. Under this technique, the respondents are asked to indicate in respect of each item whether they agree or disagree with it, and if these items form a unidimensional scale, the response pattern will be as under:

Table 5.3: Response Pattern in Scalogram Analysis

Item Number				*Respondent Score*
4	*3*	*2*	*1*	
X	X	X	X	4
–	X	X	X	3
–	–	X	X	2
–	–	–	X	1
–	–	–	–	0

X = Agree
– = Disagree

A score of 4 means that the respondent is in agreement with all the statements which is indicative of the most favourable attitude. But a score of 3 would mean that the respondent is not agreeable to item 4, but he agrees with all others. In the same way one can interpret other values of the respondents' scores. This pattern reveals that the universe of content is scalable.

Procedure: The procedure for developing a scalogram can be outlined as under:

(a) The universe of content must be defined first of all. In other words, we must lay down in clear terms the issue we want to deal within our study.

(b) The next step is to develop a number of items relating the issue and to eliminate by inspection the items that are ambiguous, irrelevant or those that happen to be too extreme items.

(c) The third step consists in pre-testing the items to determine whether the issue at hand is scalable (The pretest, as suggested by Guttman, should include 12 or more items, while the final scale may have only 4 to 6 items. Similarly, the number of respondents in a pretest may be small, say 20 or 25 but final scale should involve relatively more respondents, say 100 or more).

In a pretest the respondents are asked to record their opinions on all selected items using a Likert-type 5-point scale, ranging from 'strongly agree' to 'strongly disagree'. The strongest favourable response is scored as 5, whereas the strongest unfavourable response as 1. The total score can thus range, if there are 15 items in all, from 75 for most favourable to 15 for the least favourable.

Respondent opinionnaires are then arrayed according to total score for analysis and evaluation. If the responses of an item form a cumulative scale, its response category scores should decrease in an orderly fashion as indicated in the above table. Failure to show the said decreasing pattern means that there is overlapping which shows that the item concerned is not a good cumulative scale item i.e., the item has more than one meaning. Sometimes the overlapping in category responses can be reduced by combining categories. After analysing the pretest results, a few items, say 5 items, may be chosen.

(d) The next step is again to total the scores for the various opinionnaires, and to rearray them to reflect any shift in order, resulting from reducing the items, say, from 15 in pretest to, say, 5 for the final scale. The final pretest results may be tabulated in the form of a table given in Table 5.4.

Table 5.4: The Final Pretest Results in a Scalogram Analysis*

Scale type	*Item*					*Errors per case*	*Number of cases*	*Number of errors*
	5	*12*	*3*	*10*	*7*			
5 (perfect)	X	X	X	X	X	0	7	0
4 (perfect)	–	X	X	X	X	0	3	0
(nonscale)	–	X	–	X	X	1	1	1
(nonscale)	–	X	X	–	X	1	2	2
3 (perfect)	–	–	X	X	X	0	5	0
2 (perfect)	–	–	–	X	X	0	2	0
1 (perfect)	–	–	–	–	X	0	1	0
(nonscale)	–	–	X	–	–	2	1	2
(nonscale)	–	–	X	–	–	2	1	2
0 (perfect)	–	–	–	–	–	0	2	0
			n = 5				*N* = 25	*e* = 7

*(Figures in the table are arbitrary and have been used to explain the tabulation process only.)

The table shows that five items (numbering 5, 12, 3, 10 and 7) have been selected for the final scale. The number of respondents is 25 whose responses to various items have been tabulated along with the number of errors. Perfect scale types are those in which the respondent's answers fit the pattern that would be reproduced by using the person's total score as a guide. *Non-scale types* are those in which the category pattern differs from that expected from the respondent's total score i.e., non-scale cases have deviations from unidimensionality or errors. Whether the items (or series of statements) selected for final scale may be regarded a perfect cumulative (or a unidimensional scale), we have to examine on the basis of the coefficient of reproducibility. Guttman has set 0.9 as the level of minimum reproducibility in order to say that the scale meets the test of unidimensionality. He has given the following formula for measuring the level of reproducibility:

Guttman's Coefficient of Reproducibility = $1 - e/n(N)$

where e = number of errors

n = number of items

N = number of cases

For the above table figures,

Coefficient of Reproducibility = $1 - 7/5(25) = .94$

This shows that items number 5, 12, 3, 10 and 7 in this order constitute the cumulative or unidimensional scale, and with this we can reproduce the responses to each item, knowing only the total score of the respondent concerned.

Scalogram, analysis, like any other scaling technique, has several advantages as well as limitations. One advantage is that it assures that only a single dimension of attitude is being measured. Researcher's subjective judgement is not allowed to creep in the development of scale since the scale is determined by the replies of respondents. Then, we require only a small number of items that make such a scale easy to administer. Scalogram analysis can appropriately be used for personal, telephone or mail surveys. The main difficulty in using this scaling technique is that in practice perfect cumulative or unidimensional scales are very rarely found and we have only to use its approximation testing it through coefficient of reproducibility or examining it on the basis of some other criteria. This method is not a frequently used method for the simple reason that its development procedure is tedious and complex. Such scales hardly constitute a reliable basis for assessing attitudes of persons towards complex objects for predicting the behavioural responses of individuals towards such objects. Conceptually, this analysis is a bit more difficult in comparison to other scaling methods.

Factor Scales[*]

Factor scales are developed through factor analysis or on the basis of intercorrelations of items which indicate that a common factor accounts for the relationships between items. Factor scales are particularly "useful in uncovering latent attitude dimensions and approach scaling through the concept of multiple-dimension attribute space."[5] More specifically the two problems viz., how to deal

[*] A detailed study of the factor scales and particularly the statistical procedures involved in developing factor scales is beyond the scope of this book. As such only an introductory idea of factor scales is presented here.

[5] C. William Emory, *Business Research Methods*, p. 264-65.

appropriately with the universe of content which is multi-dimensional and how to uncover underlying (latent) dimensions which have not been identified, are dealt with through factor scales. An important factor scale based on factor analysis is *Semantic Differential* (*S.D.*) and the other one is *Multidimensional Scaling*. We give below a brief account of these factor scales.

Semantic differential scale: Semantic differential scale or the S.D. scale developed by Charles E. Osgood, G.J. Suci and P.H. Tannenbaum (1957), is an attempt to measure the psychological meanings of an object to an individual. This scale is based on the presumption that an object can have different dimensions of connotative meanings which can be located in multidimensional property space, or what can be called the semantic space in the context of S.D. scale. This scaling consists of a set of bipolar rating scales, usually of 7 points, by which one or more respondents rate one or more concepts on each scale item. For instance, the S.D. scale items for analysing candidates for leadership position may be shown as under:

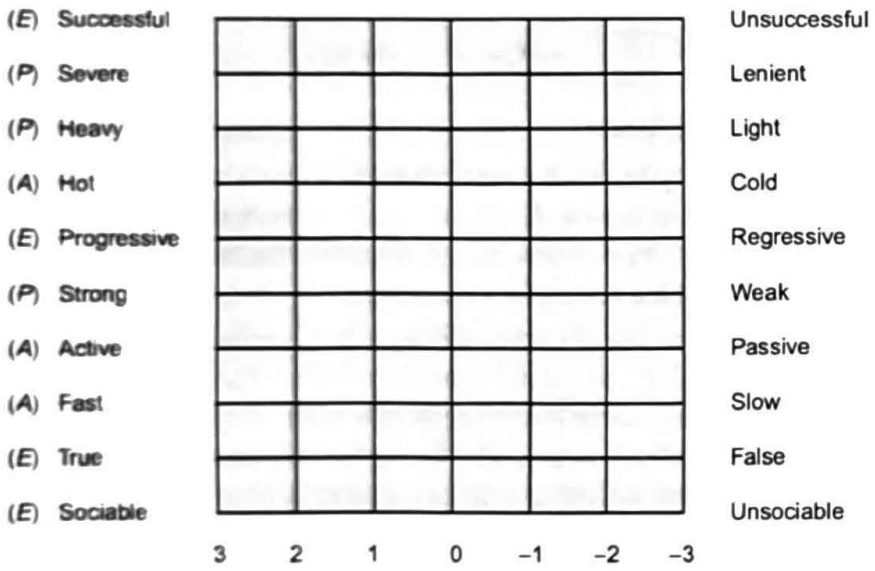


Fig. 5.4

Candidates for leadership position (along with the concept—the 'ideal' candidate) may be compared and we may score them from +3 to –3 on the basis of the above stated scales. (The letters, *E*, *P*, *A* showing the relevant factor viz., evaluation, potency and activity respectively, written along the left side are not written in actual scale. Similarly the numeric values shown are also not written in actual scale.)

Osgood and others did produce a list of some adjective pairs for attitude research purposes and concluded that semantic space is multidimensional rather than unidimensional. They made sincere efforts and ultimately found that three factors, viz., evaluation, potency and activity, contributed most to meaningful judgements by respondents. The evaluation dimension generally accounts for 1/2 and 3/4 of the extractable variance and the other two factors account for the balance.

Procedure: Various steps involved in developing S.D. scale are as follows:

(a) First of all the concepts to be studied are selected. The concepts are usually chosen by personal judgement, keeping in view the nature of the problem.

(b) The next step is to select the scales bearing in mind the criterion of factor composition and the criterion of scale's relevance to the concepts being judged (it is common practice to use at least three scales for each factor with the help of which an average factor score has to be worked out). One more criterion to be kept in view is that scales should be stable across subjects and concepts.

(c) Then a panel of judges are used to rate the various stimuli (or objects) on the various selected scales and the responses of all judges would then be combined to determine the composite scaling.

To conclude, "the S.D. has a number of specific advantages. It is an efficient and easy way to secure attitudes from a large sample. These attitudes may be measured in both direction and intensity. The total set of responses provides a comprehensive picture of the meaning of an object, as well as a measure of the subject doing the rating. It is a standardised technique that is easily repeated, but escapes many of the problems of response distortion found with more direct methods."[6]

Multidimensional scaling: Multidimensional scaling (MDS) is relatively more complicated scaling device, but with this sort of scaling one can scale objects, individuals or both with a minimum of information. Multidimensional scaling (or MDS) can be characterized as a set of procedures for portraying perceptual or affective dimensions of substantive interest. It "provides useful methodology for portraying subjective judgements of diverse kinds."[7] MDS is used when all the variables (whether metric or non-metric) in a study are to be analyzed simultaneously and all such variables happen to be independent. The underlying assumption in MDS is that people (respondents) "perceive a set of objects as being more or less similar to one another on a number of dimensions (usually uncorrelated with one another) instead of only one."[8] Through MDS techniques one can represent geometrically the locations and interrelationships among a set of points. In fact, these techniques attempt to locate the points, given the information about a set of interpoint distances, in space of one or more dimensions such as to best summarise the information contained in the interpoint distances. The distances in the solution space then optimally reflect the distances contained in the input data. For instance, if objects, say X and Y, are thought of by the respondent as being most similar as compared to all other possible pairs of objects, MDS techniques will position objects X and Y in such a way that the distance between them in multidimensional space is shorter than that between any two other objects.

Two approaches, viz., the metric approach and the non-metric approach, are usually talked about in the context of MDS, while attempting to construct a space containing m points such that $m(m-1)/2$ interpoint distances reflect the input data. The *metric approach to MDS* treats the input data as interval scale data and solves applying statistical methods for the additive constant* which

[6] *Ibid.*, p. 260.

[7] Paul E. Green, "*Analyzing Multivariate Data*", p. 421.

[8] Jagdish N. Sheth, "The Multivariate Revolution in Marketing Research", quoted in "*Marketing Research*" by Danny N. Bellenger and Barnett A. Greenberg, p. 255.

* Additive constant refers to that constant with which one can, either by subtracting or adding, convert interval scale to a ratio scale. For instance, suppose we know that distances, say a—b, b—c, c—d among stimuli on a ratio scale are 7, 6 and 3 respectively. If one were to subtract 3 from each of these distances, they would be 4, 3 and 0 respectively. The converted distances would be on an interval scale of measurement, but not on a ratio scale. Obviously, one can add 3 to all the converted distances and reachieve the ratio scale of distances. Thus 3 will be taken as the additive constant in this case. Well defined iterative approach is employed in practice for estimating appropriate additive constant.

minimises the dimensionality of the solution space. This approach utilises all the information in the data in obtaining a solution. The data (i.e., the metric similarities of the objects) are often obtained on a bipolar similarity scale on which pairs of objects are rated one at a time. If the data reflect exact distances between real objects in an *r*-dimensional space, their solution will reproduce the set of interpoint distances. But as the true and real data are rarely available, we require random and systematic procedures for obtaining a solution. Generally, the judged similarities among a set of objects are statistically transformed into distances by placing those objects in a multidimensional space of some dimensionality.

The *non-metric approach* first gathers the non-metric similarities by asking respondents to rank order all possible pairs that can be obtained from a set of objects. Such non-metric data is then transformed into some arbitrary metric space and then the solution is obtained by reducing the dimensionality. In other words, this non-metric approach seeks "a representation of points in a space of minimum dimensionality such that the rank order of the interpoint distances in the solution space maximally corresponds to that of the data. This is achieved by requiring only that the distances in the solution be monotone with the input data."[9] The non-metric approach has come into prominence during the sixties with the coming into existence of high speed computers to generate metric solutions for ordinal input data.

The significance of MDS lies in the fact that it enables the researcher to study "the perceptual structure of a set of stimuli and the cognitive processes underlying the development of this structure. Psychologists, for example, employ multidimensional scaling techniques in an effort to scale psychophysical stimuli and to determine appropriate labels for the dimensions along which these stimuli vary."[10] The MDS techniques, infact, do away with the need in the data collection process to specify the attribute(s) along which the several brands, say of a particular product, may be compared as ultimately the MDS analysis itself reveals such attribute(s) that presumably underlie the expressed relative similarities among objects. Thus, MDS is an important tool in attitude measurement and the techniques falling under MDS promise "a great advance from a series of unidimensional measurements (e.g., a distribution of intensities of feeling towards single attribute such as colour, taste or a preference ranking with indeterminate intervals), to a perceptual mapping in multidimensional space of objects ... company images, advertisement brands, etc."[11]

In spite of all the merits stated above, the MDS is not widely used because of the computation complications involved under it. Many of its methods are quite laborious in terms of both the collection of data and the subsequent analyses. However, some progress has been achieved (due to the pioneering efforts of Paul Green and his associates) during the last few years in the use of non-metric MDS in the context of market research problems. The techniques have been specifically applied in "finding out the perceptual dimensions, and the spacing of stimuli along these dimensions, that people, use in making judgements about the relative similarity of pairs of Stimuli."[12] But, "in the long run, the worth of MDS will be determined by the extent to which it advances the behavioral sciences."[13]

[9] Robert Ferber (ed.), *Handbook of Marketing Research*, p. 3–51.

[10] *Ibid.*, p. 3–52.

[11] G.B. Giles, *Marketing*, p. 43.

[12] Paul E. Green, *Analyzing Multivariate Data*, p. 421.

[13] Jum C. Nunnally, *Psychometric Theory*, p. 496.

Questions

1. What is the meaning of measurement in research? What difference does it make whether we measure in terms of a nominal, ordinal, interval or ratio scale? Explain giving examples.
2. Are you in agreement with the following statements? If so, give reasons:
   (1) Validity is more critical to measurement than reliability.
   (2) Stability and equivalence aspects of reliability essentially mean the same thing.
   (3) Content validity is the most difficult type of validity to determine.
   (4) There is no difference between concept development and concept specification.
   (5) Reliable measurement is necessarily a valid measurement.
3. Point out the possible sources of error in measurement. Describe the tests of sound measurement.
4. Are the following nominal, ordinal, interval or ratio data? Explain your answers.
   (a) Temperatures measured on the Kelvin scale.
   (b) Military ranks.
   (c) Social security numbers.
   (d) Number of passengers on buses from Delhi to Mumbai.
   (e) Code numbers given to the religion of persons attempting suicide.
5. Discuss the relative merits and demerits of:
   (a) Rating vs. Ranking scales.
   (b) Summated vs. Cumulative scales.
   (c) Scalogram analysis *vs.* Factor analysis.
6. The following table shows the results of a paired-comparison preference test of four cold drinks from a sample of 200 persons:

Name	*Coca Cola*	*Limca*	*Goldspot*	*Thumps up*
Coca Cola	–	60*	105	45
Limca	160	–	150	70
Goldspot	75	40	–	65
Thumps up	165	120	145	–

*To be read as 60 persons preferred Limca over Coca Cola.

   (a) How do these brands rank in overall preference in the given sample.
   (b) Develop an interval scale for the four varieties of cold drinks.
7. (1) Narrate the procedure for developing a scalogram and illustrate the same by an example.
   (2) Workout Guttman's coefficient of reproducibility from the following information:
   Number of cases (N) = 30
   Number of items (n) = 6
   Number of errors (e) = 10
   Interpret the meaning of coefficient you work out in this example.
8. Write short notes on:
   (a) Semantic differential scale;
   (b) Scalogram analysis;

(c) Likert-type scale;

(d) Arbitrary scales;

(e) Multidimensional scaling (MDS).

9. Describe the different methods of scale construction, pointing out the merits and demerits of each.

10. "Scaling describes the procedures by which numbers are assigned to various degrees of opinion, attitude and other concepts." Discuss. Also point out the bases for scale classification.

6

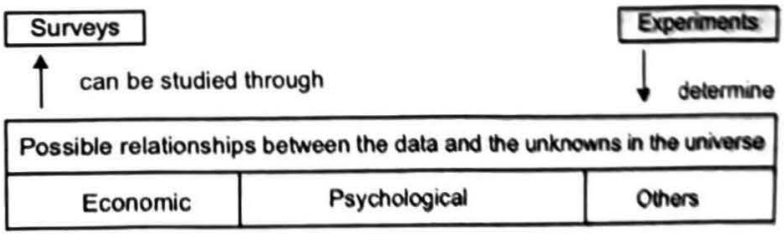
Methods of Data Collection

The task of data collection begins after a research problem has been defined and research design/ plan chalked out. While deciding about the method of data collection to be used for the study, the researcher should keep in mind two types of data viz., primary and secondary. The *primary data* are those which are collected afresh and for the first time, and thus happen to be original in character. The *secondary data,* on the other hand, are those which have already been collected by someone else and which have already been passed through the statistical process. The researcher would have to decide which sort of data he would be using (thus collecting) for his study and accordingly he will have to select one or the other method of data collection. The methods of collecting primary and secondary data differ since primary data are to be originally collected, while in case of secondary data the nature of data collection work is merely that of compilation. We describe the different methods of data collection, with the pros and cons of each method.

COLLECTION OF PRIMARY DATA

We collect primary data during the course of doing experiments in an experimental research but in case we do research of the descriptive type and perform surveys, whether sample surveys or census surveys, then we can obtain primary data either through observation or through direct communication with respondents in one form or another or through personal interviews.* This, in other words, means

*An experiment refers to an investigation in which a factor or variable under test is isolated and its effect(s) measured. In an experiment the investigator measures the effects of an experiment which he conducts intentionally. Survey refers to the method of securing information concerning a phenomena under study from all or a selected number of respondents of the concerned universe. In a survey, the investigator examines those phenomena which exist in the universe independent of his action. The difference between an experiment and a survey can be depicted as under:



that there are several methods of collecting primary data, particularly in surveys and descriptive researches. Important ones are: (i) observation method, (ii) interview method, (iii) through questionnaires, (iv) through schedules, and (v) other methods which include (a) warranty cards; (b) distributor audits; (c) pantry audits; (d) consumer panels; (e) using mechanical devices; (f) through projective techniques; (g) depth interviews, and (h) content analysis. We briefly take up each method separately.

Observation Method

The observation method is the most commonly used method specially in studies relating to behavioural sciences. In a way we all observe things around us, but this sort of observation is not scientific observation. Observation becomes a scientific tool and the method of data collection for the researcher, when it serves a formulated research purpose, is systematically planned and recorded and is subjected to checks and controls on validity and reliability. Under the observation method, the information is sought by way of investigator's own direct observation without asking from the respondent. For instance, in a study relating to consumer behaviour, the investigator instead of asking the brand of wrist watch used by the respondent, may himself look at the watch. The main advantage of this method is that subjective bias is eliminated, if observation is done accurately. Secondly, the information obtained under this method relates to what is currently happening; it is not complicated by either the past behaviour or future intentions or attitudes. Thirdly, this method is independent of respondents' willingness to respond and as such is relatively less demanding of active cooperation on the part of respondents as happens to be the case in the interview or the questionnaire method. This method is particularly suitable in studies which deal with subjects (i.e., respondents) who are not capable of giving verbal reports of their feelings for one reason or the other

However, observation method has various limitations. Firstly, it is an expensive method. Secondly, the information provided by this method is very limited. Thirdly, sometimes unforeseen factors may interfere with the observational task. At times, the fact that some people are rarely accessible to direct observation creates obstacle for this method to collect data effectively.

While using this method, the researcher should keep in mind things like: What should be observed? How the observations should be recorded? Or how the accuracy of observation can be ensured? In case the observation is characterised by a careful definition of the units to be observed, the style of recording the observed information, standardised conditions of observation and the selection of pertinent data of observation, then the observation is called as *structured observation.* But when observation is to take place without these characteristics to be thought of in advance, the same is termed *as unstructured observation.* Structured observation is considered appropriate in descriptive studies, whereas in an exploratory study the observational procedure is most likely to be relatively unstructured.

We often talk about participant and non-participant types of observation in the context of studies, particularly of social sciences. This distinction depends upon the observer's sharing or not sharing the life of the group he is observing. If the observer observes by making himself, more or less, a member of the group he is observing so that he can experience what the members of the group experience, the observation is called as the *participant observation.* But when the observer observes as a detached emissary without any attempt on his part to experience through participation what others feel, the observation of this type is often termed as *non-participant observation.* (When the observer is observing in such a manner that his presence may be unknown to the people he is observing, such an observation is described as *disguised observation.*)

There are several merits of the participant type of observation: (i) The researcher is enabled to record the natural behaviour of the group. (ii) The researcher can even gather information which could not easily be obtained if he observes in a disinterested fashion. (iii) The researcher can even verify the truth of statements made by informants in the context of a questionnaire or a schedule. But there are also certain demerits of this type of observation viz., the observer may lose the objectivity to the extent he participates emotionally; the problem of observation-control is not solved; and it may narrow-down the researcher's range of experience.

Sometimes we talk of *controlled* and *uncontrolled observation.* If the observation takes place in the natural setting, it may be termed as uncontrolled observation, but when observation takes place according to definite pre-arranged plans, involving experimental procedure, the same is then termed controlled observation. In non-controlled observation, no attempt is made to use precision instruments. The major aim of this type of observation is to get a spontaneous picture of life and persons. It has a tendency to supply naturalness and completeness of behaviour, allowing sufficient time for observing it. But in controlled observation, we use mechanical (or precision) instruments as aids to accuracy and standardisation. Such observation has a tendency to supply formalised data upon which generalisations can be built with some degree of assurance. The main pitfall of non-controlled observation is that of subjective interpretation. There is also the danger of having the feeling that we know more about the observed phenomena than we actually do. Generally, controlled observation takes place in various experiments that are carried out in a laboratory or under controlled conditions, whereas uncontrolled observation is resorted to in case of exploratory researches.

Interview Method

The interview method of collecting data involves presentation of oral-verbal stimuli and reply in terms of oral-verbal responses. This method can be used through personal interviews and, if possible, through telephone interviews.

(a) *Personal interviews:* Personal interview method requires a person known as the interviewer asking questions generally in a face-to-face contact to the other person or persons. (At times the interviewee may also ask certain questions and the interviewer responds to these, but usually the interviewer initiates the interview and collects the information.) This sort of interview may be in the form of direct personal investigation or it may be indirect oral investigation. In the case of direct personal investigation the interviewer has to collect the information personally from the sources concerned. He has to be on the spot and has to meet people from whom data have to be collected. This method is particularly suitable for intensive investigations. But in certain cases it may not be possible or worthwhile to contact directly the persons concerned or on account of the extensive scope of enquiry, the direct personal investigation technique may not be used. In such cases an indirect oral examination can be conducted under which the interviewer has to cross-examine other persons who are supposed to have knowledge about the problem under investigation and the information, obtained is recorded. Most of the commissions and committees appointed by government to carry on investigations make use of this method.

The method of collecting information through personal interviews is usually carried out in a structured way. As such we call the interviews as *structured interviews.* Such interviews involve the use of a set of predetermined questions and of highly standardised techniques of recording. Thus,

the interviewer in a structured interview follows a rigid procedure laid down, asking questions in a form and order prescribed. As against it, the *unstructured interviews* are characterised by a flexibility of approach to questioning. Unstructured interviews do not follow a system of pre-determined questions and standardised techniques of recording information. In a non-structured interview, the interviewer is allowed much greater freedom to ask, in case of need, supplementary questions or at times he may omit certain questions if the situation so requires. He may even change the sequence of questions. He has relatively greater freedom while recording the responses to include some aspects and exclude others. But this sort of flexibility results in lack of comparability of one interview with another and the analysis of unstructured responses becomes much more difficult and time-consuming than that of the structured responses obtained in case of structured interviews. Unstructured interviews also demand deep knowledge and greater skill on the part of the interviewer. Unstructured interview, however, happens to be the central technique of collecting information in case of exploratory or formulative research studies. But in case of descriptive studies, we quite often use the technique of structured interview because of its being more economical, providing a safe basis for generalisation and requiring relatively lesser skill on the part of the interviewer.

We may as well talk about focussed interview, clinical interview and the non-directive interview. *Focussed interview* is meant to focus attention on the given experience of the respondent and its effects. Under it the interviewer has the freedom to decide the manner and sequence in which the questions would be asked and has also the freedom to explore reasons and motives. The main task of the interviewer in case of a focussed interview is to confine the respondent to a discussion of issues with which he seeks conversance. Such interviews are used generally in the development of hypotheses and constitute a major type of unstructured interviews. The *clinical interview* is concerned with broad underlying feelings or motivations or with the course of individual's life experience. The method of eliciting information under it is generally left to the interviewer's discretion. In case of *non-directive interview*, the interviewer's function is simply to encourage the respondent to talk about the given topic with a bare minimum of direct questioning. The interviewer often acts as a catalyst to a comprehensive expression of the respondents' feelings and beliefs and of the frame of reference within which such feelings and beliefs take on personal significance.

Despite the variations in interview-techniques, the major advantages and weaknesses of personal interviews can be enumerated in a general way. The chief merits of the interview method are as follows

(i) More information and that too in greater depth can be obtained.

(ii) Interviewer by his own skill can overcome the resistance, if any, of the respondents; the interview method can be made to yield an almost perfect sample of the general population.

(iii) There is greater flexibility under this method as the opportunity to restructure questions is always there, specially in case of unstructured interviews.

(iv) Observation method can as well be applied to recording verbal answers to various questions.

(v) Personal information can as well be obtained easily under this method.

(vi) Samples can be controlled more effectively as there arises no difficulty of the missing returns, non-response generally remains very low.

(vii) The interviewer can usually control which person(s) will answer the questions. This is not possible in mailed questionnaire approach. If so desired, group discussions may also be held

(viii) The interviewer may catch the informant off-guard and thus may secure the most spontaneous reactions than would be the case if mailed questionnaire is used.

(ix) The language of the interview can be adopted to the ability or educational level of the person interviewed and as such misinterpretations concerning questions can be avoided.

(x) The interviewer can collect supplementary information about the respondent's personal characteristics and environment which is often of great value in interpreting results.

But there are also certain weaknesses of the interview method. Among the important weaknesses, mention may be made of the following:

(i) It is a very expensive method, specially when large and widely spread geographical sample is taken.

(ii) There remains the possibility of the bias of interviewer as well as that of the respondent; there also remains the headache of supervision and control of interviewers.

(iii) Certain types of respondents such as important officials or executives or people in high income groups may not be easily approachable under this method and to that extent the data may prove inadequate.

(iv) This method is relatively more-time-consuming, specially when the sample is large and re-calls upon the respondents are necessary.

(v) The presence of the interviewer on the spot may over-stimulate the respondent, sometimes even to the extent that he may give imaginary information just to make the interview interesting.

(vi) Under the interview method the organisation required for selecting, training and supervising the field-staff is more complex with formidable problems.

(vii) Interviewing at times may also introduce systematic errors.

(viii) Effective interview presupposes proper rapport with respondents that would facilitate free and frank responses. This is often a very difficult requirement.

Pre-requisites and basic tenets of interviewing: For successful implementation of the interview method, interviewers should be carefully selected, trained and briefed. They should be honest, sincere, hardworking, impartial and must possess the technical competence and necessary practical experience. Occasional field checks should be made to ensure that interviewers are neither cheating, nor deviating from instructions given to them for performing their job efficiently. In addition, some provision should also be made in advance so that appropriate action may be taken if some of the selected respondents refuse to cooperate or are not available when an interviewer calls upon them.

In fact, interviewing is an art governed by certain scientific principles. Every effort should be made to create friendly atmosphere of trust and confidence, so that respondents may feel at ease while talking to and discussing with the interviewer. The interviewer must ask questions properly and intelligently and must record the responses accurately and completely. At the same time, the interviewer must answer legitimate question(s), if any, asked by the respondent and must clear any doubt that the latter has. The interviewers approach must be friendly, courteous, conversational and unbiased. The interviewer should not show surprise or disapproval of a respondent's answer but he must keep the direction of interview in his own hand, discouraging irrelevant conversation and must make all possible effort to keep the respondent on the track.

(b) *Telephone interviews:* This method of collecting information consists in contacting respondents on telephone itself. It is not a very widely used method, but plays important part in industrial surveys, particularly in developed regions. The chief merits of such a system are:

1. It is more flexible in comparison to mailing method.
2. It is faster than other methods i.e., a quick way of obtaining information.
3. It is cheaper than personal interviewing method; here the cost per response is relatively low.
4. Recall is easy; callbacks are simple and economical.
5. There is a higher rate of response than what we have in mailing method; the non-response is generally very low.
6. Replies can be recorded without causing embarrassment to respondents.
7. Interviewer can explain requirements more easily.
8. At times, access can be gained to respondents who otherwise cannot be contacted for one reason or the other.
9. No field staff is required.
10. Representative and wider distribution of sample is possible.

But this system of collecting information is not free from demerits. Some of these may be highlighted.

1. Little time is given to respondents for considered answers; interview period is not likely to exceed five minutes in most cases.
2. Surveys are restricted to respondents who have telephone facilities.
3. Extensive geographical coverage may get restricted by cost considerations.
4. It is not suitable for intensive surveys where comprehensive answers are required to various questions.
5. Possibility of the bias of the interviewer is relatively more.
6. Questions have to be short and to the point; probes are difficult to handle.

COLLECTION OF DATA THROUGH QUESTIONNAIRES

This method of data collection is quite popular, particularly in case of big enquiries. It is being adopted by private individuals, research workers, private and public organisations and even by governments. In this method a questionnaire is sent (usually by post) to the persons concerned with a request to answer the questions and return the questionnaire. A questionnaire consists of a number of questions printed or typed in a definite order on a form or set of forms. The questionnaire is mailed to respondents who are expected to read and understand the questions and write down the reply in the space meant for the purpose in the questionnaire itself. The respondents have to answer the questions on their own.

The method of collecting data by mailing the questionnaires to respondents is most extensively employed in various economic and business surveys. The merits claimed on behalf of this method are as follows:

1. There is low cost even when the universe is large and is widely spread geographically.

2. It is free from the bias of the interviewer; answers are in respondents' own words.
3. Respondents have adequate time to give well thought out answers.
4. Respondents, who are not easily approachable, can also be reached conveniently.
5 Large samples can be made use of and thus the results can be made more dependable and reliable.

The main demerits of this system can also be listed here:

1. Low rate of return of the duly filled in questionnaires; bias due to no-response is often indeterminate.
2. It can be used only when respondents are educated and cooperating.
3. The control over questionnaire may be lost once it is sent.
4. There is inbuilt inflexibility because of the difficulty of amending the approach once questionnaires have been despatched.
5. There is also the possibility of ambiguous replies or omission of replies altogether to certain questions; interpretation of omissions is difficult.
6. It is difficult to know whether willing respondents are truly representative.
7. This method is likely to be the slowest of all.

Before using this method, it is always advisable to conduct 'pilot study' (Pilot Survey) for testing the questionnaires. In a big enquiry the significance of pilot survey is felt very much. Pilot survey is infact the replica and rehearsal of the main survey. Such a survey, being conducted by experts, brings to the light the weaknesses (if any) of the questionnaires and also of the survey techniques. From the experience gained in this way, improvement can be effected.

Main aspects of a questionnaire: Quite often questionnaire is considered as the heart of a survey operation. Hence it should be very carefully constructed. If it is not properly set up, then the survey is bound to fail. This fact requires us to study the main aspects of a questionnaire viz., the general form, question sequence and question formulation and wording. Researcher should note the following with regard to these three main aspects of a questionnaire:

1. *General form:* So far as the general form of a questionnaire is concerned, it can either be structured or unstructured questionnaire. Structured questionnaires are those questionnaires in which there are definite, concrete and pre-determined questions. The questions are presented with exactly the same wording and in the same order to all respondents. Resort is taken to this sort of standardisation to ensure that all respondents reply to the same set of questions. The form of the question may be either closed (i.e., of the type 'yes' or 'no') or open (i.e., inviting free response) but should be stated in advance and not constructed during questioning. Structured questionnaires may also have fixed alternative questions in which responses of the informants are limited to the stated alternatives. Thus a highly structured questionnaire is one in which all questions and answers are specified and comments in the respondent's own words are held to the minimum. When these characteristics are not present in a questionnaire, it can be termed as unstructured or non-structured questionnaire. More specifically, we can say that in an unstructured questionnaire, the interviewer is provided with a general guide on the type of information to be obtained, but the exact question formulation is largely his own responsibility and the replies are to be taken down in the respondent's own words to the extent possible; in some situations tape recorders may be used to achieve this goal.

Structured questionnaires are simple to administer and relatively inexpensive to analyse. The provision of alternative replies, at times, helps to understand the meaning of the question clearly. But such questionnaires have limitations too. For instance, wide range of data and that too in respondent's own words cannot be obtained with structured questionnaires. They are usually considered inappropriate in investigations where the aim happens to be to probe for attitudes and reasons for certain actions or feelings. They are equally not suitable when a problem is being first explored and working hypotheses sought. In such situations, unstructured questionnaires may be used effectively. Then on the basis of the results obtained in pretest (testing before final use) operations from the use of unstructured questionnaires, one can construct a structured questionnaire for use in the main study.

2. *Question sequence:* In order to make the questionnaire effective and to ensure quality to the replies received, a researcher should pay attention to the question-sequence in preparing the questionnaire. A proper sequence of questions reduces considerably the chances of individual questions being misunderstood. The question-sequence must be clear and smoothly-moving, meaning thereby that the relation of one question to another should be readily apparent to the respondent, with questions that are easiest to answer being put in the beginning. The first few questions are particularly important because they are likely to influence the attitude of the respondent and in seeking his desired cooperation. The opening questions should be such as to arouse human interest. The following type of questions should generally be avoided as opening questions in a questionnaire:

1. questions that put too great a strain on the memory or intellect of the respondent;
2. questions of a personal character;
3. questions related to personal wealth, etc.

Following the opening questions, we should have questions that are really vital to the research problem and a connecting thread should run through successive questions. Ideally, the question-sequence should conform to the respondent's way of thinking. Knowing what information is desired, the researcher can rearrange the order of the questions (this is possible in case of unstructured questionnaire) to fit the discussion in each particular case. But in a structured questionnaire the best that can be done is to determine the question-sequence with the help of a Pilot Survey which is likely to produce good rapport with most respondents. Relatively difficult questions must be relegated towards the end so that even if the respondent decides not to answer such questions, considerable information would have already been obtained. Thus, question-sequence should usually go from the general to the more specific and the researcher must always remember that the answer to a given question is a function not only of the question itself, but of all previous questions as well. For instance, if one question deals with the price usually paid for coffee and the next with reason for preferring that particular brand, the answer to this latter question may be couched largely in terms of price-differences.

3. *Question formulation and wording:* With regard to this aspect of questionnaire, the researcher should note that each question must be very clear for any sort of misunderstanding can do irreparable harm to a survey. Question should also be impartial in order not to give a biased picture of the true state of affairs. Questions should be constructed with a view to their forming a logical part of a well thought out tabulation plan. In general, all questions should meet the following standards—(a) should be easily understood; (b) should be simple i.e., should convey only one thought at a time; (c) should be concrete and should conform as much as possible to the respondent's way of thinking. (For

instance, instead of asking. "How many razor blades do you use annually?" The more realistic question would be to ask, "How many razor blades did you use last week?"

Concerning the form of questions, we can talk about two principal forms, viz., multiple choice question and the open-end question. In the former the respondent selects one of the alternative possible answers put to him, whereas in the latter he has to supply the answer in his own words. The question with only two possible answers (usually 'Yes' or 'No') can be taken as a special case of the multiple choice question, or can be named as a 'closed question.' There are some advantages and disadvantages of each possible form of question. Multiple choice or closed questions have the advantages of easy handling, simple to answer, quick and relatively inexpensive to analyse. They are most amenable to statistical analysis. Sometimes, the provision of alternative replies helps to make clear the meaning of the question. But the main drawback of fixed alternative questions is that of "putting answers in people's mouths" i.e., they may force a statement of opinion on an issue about which the respondent does not infact have any opinion. They are not appropriate when the issue under consideration happens to be a complex one and also when the interest of the researcher is in the exploration of a process. In such situations, open-ended questions which are designed to permit a free response from the respondent rather than one limited to certain stated alternatives are considered appropriate. Such questions give the respondent considerable latitude in phrasing a reply. Getting the replies in respondent's own words is, thus, the major advantage of open-ended questions. But one should not forget that, from an analytical point of view, open-ended questions are more difficult to handle, raising problems of interpretation, comparability and interviewer bias.*

In practice, one rarely comes across a case when one questionnaire relies on one form of questions alone. The various forms complement each other. As such questions of different forms are included in one single questionnaire. For instance, multiple-choice questions constitute the basis of a structured questionnaire, particularly in a mail survey. But even there, various open-ended questions are generally inserted to provide a more complete picture of the respondent's feelings and attitudes.

Researcher must pay proper attention to the wordings of questions since reliable and meaningful returns depend on it to a large extent. Since words are likely to affect responses, they should be properly chosen. Simple words, which are familiar to all respondents should be employed. Words with ambiguous meanings must be avoided. Similarly, danger words, catch-words or words with emotional connotations should be avoided. Caution must also be exercised in the use of phrases which reflect upon the prestige of the respondent. Question wording, in no case, should bias the answer. In fact, question wording and formulation is an art and can only be learnt by practice.

Essentials of a good questionnaire: To be successful, questionnaire should be comparatively short and simple i.e., the size of the questionnaire should be kept to the minimum. Questions should proceed in logical sequence moving from easy to more difficult questions. Personal and intimate questions should be left to the end. Technical terms and vague expressions capable of different interpretations should be avoided in a questionnaire. Questions may be dichotomous (yes or no answers), multiple choice (alternative answers listed) or open-ended. The latter type of questions are often difficult to analyse and hence should be avoided in a questionnaire to the extent possible. There should be some control questions in the questionnaire which indicate the reliability of the respondent. For instance, a question designed to determine the consumption of particular material may be asked

* Interviewer bias refers to the extent to which an answer is altered in meaning by some action or attitude on the part of the interviewer.

first in terms of financial expenditure and later in terms of weight. The control questions, thus, introduce a cross-check to see whether the information collected is correct or not. Questions affecting the sentiments of respondents should be avoided. Adequate space for answers should be provided in the questionnaire to help editing and tabulation. There should always be provision for indications of uncertainty, e.g., "do not know," "no preference" and so on. Brief directions with regard to filling up the questionnaire should invariably be given in the questionnaire itself. Finally, the physical appearance of the questionnaire affects the cooperation the researcher receives from the recipients and as such an attractive looking questionnaire, particularly in mail surveys, is a plus point for enlisting cooperation. The quality of the paper, along with its colour, must be good so that it may attract the attention of recipients.

COLLECTION OF DATA THROUGH SCHEDULES

This method of data collection is very much like the collection of data through questionnaire, with little difference which lies in the fact that schedules (proforma containing a set of questions) are being filled in by the enumerators who are specially appointed for the purpose. These enumerators along with schedules, go to respondents, put to them the questions from the proforma in the order the questions are listed and record the replies in the space meant for the same in the proforma. In certain situations, schedules may be handed over to respondents and enumerators may help them in recording their answers to various questions in the said schedules. Enumerators explain the aims and objects of the investigation and also remove the difficulties which any respondent may feel in understanding the implications of a particular question or the definition or concept of difficult terms.

This method requires the selection of enumerators for filling up schedules or assisting respondents to fill up schedules and as such enumerators should be very carefully selected. The enumerators should be trained to perform their job well and the nature and scope of the investigation should be explained to them thoroughly so that they may well understand the implications of different questions put in the schedule. Enumerators should be intelligent and must possess the capacity of cross-examination in order to find out the truth. Above all, they should be honest, sincere, hardworking and should have patience and perseverance.

This method of data collection is very useful in extensive enquiries and can lead to fairly reliable results. It is, however, very expensive and is usually adopted in investigations conducted by governmental agencies or by some big organisations. Population census all over the world is conducted through this method.

DIFFERENCE BETWEEN QUESTIONNAIRES AND SCHEDULES

Both questionnaire and schedule are popularly used methods of collecting data in research surveys. There is much resemblance in the nature of these two methods and this fact has made many people to remark that from a practical point of view, the two methods can be taken to be the same. But from the technical point of view there is difference between the two. The important points of difference are as under:

1. The questionnaire is generally sent through mail to informants to be answered as specified in a covering letter, but otherwise without further assistance from the sender. The schedule

is generally filled out by the research worker or the enumerator, who can interpret questions when necessary.

2. To collect data through questionnaire is relatively cheap and economical since we have to spend money only in preparing the questionnaire and in mailing the same to respondents. Here no field staff required. To collect data through schedules is relatively more expensive since considerable amount of money has to be spent in appointing enumerators and in importing training to them. Money is also spent in preparing schedules.
3. Non-response is usually high in case of questionnaire as many people do not respond and many return the questionnaire without answering all questions. Bias due to non-response often remains indeterminate. As against this, non-response is generally very low in case of schedules because these are filled by enumerators who are able to get answers to all questions. But there remains the danger of interviewer bias and cheating.
4. In case of questionnaire, it is not always clear as to who replies, but in case of schedule the identity of respondent is known.
5. The questionnaire method is likely to be very slow since many respondents do not return the questionnaire in time despite several reminders, but in case of schedules the information is collected well in time as they are filled in by enumerators.
6. Personal contact is generally not possible in case of the questionnaire method as questionnaires are sent to respondents by post who also in turn return the same by post. But in case of schedules direct personal contact is established with respondents.
7. Questionnaire method can be used only when respondents are literate and cooperative, but in case of schedules the information can be gathered even when the respondents happen to be illiterate.
8. Wider and more representative distribution of sample is possible under the questionnaire method, but in respect of schedules there usually remains the difficulty in sending enumerators over a relatively wider area.
9. Risk of collecting incomplete and wrong information is relatively more under the questionnaire method, particularly when people are unable to understand questions properly. But in case of schedules, the information collected is generally complete and accurate as enumerators can remove the difficulties, if any, faced by respondents in correctly understanding the questions. As a result, the information collected through schedules is relatively more accurate than that obtained through questionnaires.
10. The success of questionnaire method lies more on the quality of the questionnaire itself, but in the case of schedules much depends upon the honesty and competence of enumerators.
11. In order to attract the attention of respondents, the physical appearance of questionnaire must be quite attractive, but this may not be so in case of schedules as they are to be filled in by enumerators and not by respondents.
12. Along with schedules, observation method can also be used but such a thing is not possible while collecting data through questionnaires.

SOME OTHER METHODS OF DATA COLLECTION

Let us consider some other methods of data collection, particularly used by big business houses in modern times.

1. Warranty cards: Warranty cards are usually postal sized cards which are used by dealers of consumer durables to collect information regarding their products. The information sought is printed in the form of questions on the 'warranty cards' which is placed inside the package along with the product with a request to the consumer to fill in the card and post it back to the dealer.

2. Distributor or store audits: Distributor or store audits are performed by distributors as well as manufactures through their salesmen at regular intervals. Distributors get the retail stores audited through salesmen and use such information to estimate market size, market share, seasonal purchasing pattern and so on. The data are obtained in such audits not by questioning but by observation. For instance, in case of a grocery store audit, a sample of stores is visited periodically and data are recorded on inventories on hand either by observation or copying from store records. Store audits are invariably panel operations, for the derivation of sales estimates and compilation of sales trends by stores are their principal '*raison detre*'. The principal advantage of this method is that it offers the most efficient way of evaluating the effect on sales of variations of different techniques of in-store promotion.

3. Pantry audits: Pantry audit technique is used to estimate consumption of the basket of goods at the consumer level. In this type of audit, the investigator collects an inventory of types, quantities and prices of commodities consumed. Thus in pantry audit data are recorded from the examination of consumer's pantry. The usual objective in a pantry audit is to find out what types of consumers buy certain products and certain brands, the assumption being that the contents of the pantry accurately portray consumer's preferences. Quite often, pantry audits are supplemented by direct questioning relating to reasons and circumstances under which particular products were purchased in an attempt to relate these factors to purchasing habits. A pantry audit may or may not be set up as a panel operation, since a single visit is often considered sufficient to yield an accurate picture of consumers' preferences. An important limitation of pantry audit approach is that, at times, it may not be possible to identify consumers' preferences from the audit data alone, particularly when promotion devices produce a marked rise in sales.

4. Consumer panels: An extension of the pantry audit approach on a regular basis is known as 'consumer panel', where a set of consumers are arranged to come to an understanding to maintain detailed daily records of their consumption and the same is made available to investigator on demands. In other words, a consumer panel is essentially a sample of consumers who are interviewed repeatedly over a period of time. Mostly consume panels are of two types viz., the transitory consumer panel and the continuing consumer panel. *A transitory consumer panel* is set up to measure the effect of a particular phenomenon. Usually such a panel is conducted on a before-and-after-basis. Initial interviews are conducted before the phenomenon takes place to record the attitude of the consumer. A second set of interviews is carried out after the phenomenon has taken place to find out the consequent changes that might have occurred in the consumer's attitude. It is a favourite tool of advertising and of social research. A *continuing consumer panel* is often set up for an indefinite period with a view to collect data on a particular aspect of consumer behaviour over time, generally at periodic intervals or may be meant to serve as a general purpose panel for researchers on a variety of subjects. Such panels have been used in the area of consumer expenditure, public opinion and radio and TV listenership

among others. Most of these panels operate by mail. The representativeness of the panel relative to the population and the effect of panel membership on the information obtained after the two major problems associated with the use of this method of data collection.

5. Use of mechanical devices: The use of mechanical devices has been widely made to collect information by way of indirect means. Eye camera, Pupilometric camera, Psychogalvanometer, Motion picture camera and Audiometer are the principal devices so far developed and commonly used by modern big business houses, mostly in the developed world for the purpose of collecting the required information.

Eye cameras are designed to record the focus of eyes of a respondent on a specific portion of a sketch or diagram or written material. Such an information is useful in designing advertising material. Pupilometric cameras record dilation of the pupil as a result of a visual stimulus. The extent of dilation shows the degree of interest aroused by the stimulus. Psychogalvanometer is used for measuring the extent of body excitement as a result of the visual stimulus. Motion picture cameras can be used to record movement of body of a buyer while deciding to buy a consumer good from a shop or big store. Influence of packaging or the information given on the label would stimulate a buyer to perform certain physical movements which can easily be recorded by a hidden motion picture camera in the shop's four walls. Audiometers are used by some TV concerns to find out the type of programmes as well as stations preferred by people. A device is fitted in the television instrument itself to record these changes. Such data may be used to find out the market share of competing television stations.

6. Projective techniques: Projective techniques (or what are sometimes called as indirect interviewing techniques) for the collection of data have been developed by psychologists to use projections of respondents for inferring about underlying motives, urges, or intentions which are such that the respondent either resists to reveal them or is unable to figure out himself. In projective techniques the respondent in supplying information tends unconsciously to project his own attitudes or feelings on the subject under study. Projective techniques play an important role in motivational researches or in attitude surveys.

The use of these techniques requires intensive specialised training. In such techniques, the individual's responses to the stimulus-situation are not taken at their face value. The stimuli may arouse many different kinds of reactions. The nature of the stimuli and the way in which they are presented under these techniques do not clearly indicate the way in which the response is to be interpreted. The stimulus may be a photograph, a picture, an inkblot and so on. Responses to these stimuli are interpreted as indicating the individual's own view, his personality structure, his needs, tensions, etc. in the context of some pre-established psychological conceptualisation of what the individual's responses to the stimulus mean.

We may now briefly deal with the important projective techniques.

(i) *Word association tests:* These tests are used to extract information regarding such words which have maximum association. In this sort of test the respondent is asked to mention the first word that comes to mind, ostensibly without thinking, as the interviewer reads out each word from a list. If the interviewer says *cold*, the respondent may say *hot* and the like ones. The general technique is to use a list of as many as 50 to 100 words. Analysis of the matching words supplied by the respondents indicates whether the given word should be used for the contemplated purpose. The same idea is exploited in marketing research to find out the quality that is mostly associated to a brand of a product. A number of qualities of a product may be listed and informants may be asked to write

brand names possessing one or more of these. This technique is quick and easy to use, but yields reliable results when applied to words that are widely known and which possess essentially one type of meaning. This technique is frequently used in advertising research.

(ii) *Sentence completion tests:* These tests happen to be an extension of the technique of word association tests. Under this, informant may be asked to complete a sentence (such as: persons who wear Khadi are...) to find association of Khadi clothes with certain personality characteristics. Several sentences of this type might be put to the informant on the same subject. Analysis of replies from the same informant reveals his attitude toward that subject, and the combination of these attitudes of all the sample members is then taken to reflect the views of the population. This technique permits the testing not only of words (as in case of word association tests), but of ideas as well and thus, helps in developing hypotheses and in the construction of questionnaires. This technique is also quick and easy to use, but it often leads to analytical problems, particularly when the response happens to be multidimensional.

(iii) *Story completion tests:* Such tests are a step further wherein the researcher may contrive stories instead of sentences and ask the informants to complete them. The respondent is given just enough of story to focus his attention on a given subject and he is asked to supply a conclusion to the story.

(iv) *Verbal projection tests:* These are the tests wherein the respondent is asked to comment on or to explain what other people do. For example, why do people smoke? Answers may reveal the respondent's own motivations.

(v) *Pictorial techniques:* There are several pictorial techniques. The important ones are as follows:

(a) *Thematic apperception test (T.A.T.):* The TAT consists of a set of pictures (some of the pictures deal with the ordinary day-to-day events while others may be ambiguous pictures of unusual situations) that are shown to respondents who are asked to describe what they think the pictures represent. The replies of respondents constitute the basis for the investigator to draw inferences about their personality structure, attitudes, etc.

(b) *Rosenzweig test:* This test uses a cartoon format wherein we have a series of cartoons with words inserted in 'balloons' above. The respondent is asked to put his own words in an empty balloon space provided for the purpose in the picture. From what the respondents write in this fashion, the study of their attitudes can be made.

(c) *Rorschach test:* This test consists of ten cards having prints of inkblots. The design happens to be symmetrical but meaningless. The respondents are asked to describe what they perceive in such symmetrical inkblots and the responses are interpreted on the basis of some pre-determined psychological framework. This test is frequently used but the problem of validity still remains a major problem of this test.

(d) *Holtzman Inkblot Test (HIT):* This test from W.H. Holtzman is a modification of the Rorschach Test explained above. This test consists of 45 inkblot cards (and not 10 inkblots as we find in case of Rorschach Test) which are based on colour, movement, shading and other factors involved in inkblot perception. Only one response per card is obtained from the subject (or the respondent) and the responses of a subject are interpreted at three levels of form appropriateness. Form responses are interpreted for knowing the accuracy (F) or inaccuracy (F–) of respondent's percepts; shading and colour for ascertaining his affectional and emotional needs; and movement responses for assessing the dynamic aspects of his life.

Holtzman Inkblot Test or H.I.T. has several special features or advantages. For example, it elicits relatively constant number of responses per respondent. Secondly, it facilitates studying the responses of a respondent to different cards in the light of norms of each card instead of lumping them together. Thirdly, it elicits much more information from the respondent then is possible with merely 10 cards in Rorschach test; the 45 cards used in this test provide a variety of stimuli to the respondent and as such the range of responses elicited by the test is comparatively wider.

There are some limitations of this test as well. One difficulty that remains in using this test is that most of the respondents do not know the determinants of their perceptions, but for the researcher, who has to interpret the protocols of a subject and understand his personality (or attitude) through them, knowing the determinant of each of his response is a must. This fact emphasises that the test must be administered individually and a post-test inquiry must as well be conducted for knowing the nature and sources of responses and this limits the scope of HIT as a group test of personality. Not only this, "the usefulness of HIT for purposes of personal selection, vocational guidance, etc. is still to be established."[1]

In view of these limitations, some people have made certain changes in applying this test. For instance, Fisher and Cleveland in their approach for obtaining Barrier score of an individual's personality have developed a series of multiple choice items for 40 of HIT cards. Each of these cards is presented to the subject along with three acceptable choices [such as 'Knight in armour' (Barrier response), 'X-Ray' (Penetrating response) and 'Flower' (Neutral response)]. Subject taking the test is to check the choice he likes most, make a different mark against the one he likes least and leave the third choice blank. The number of barrier responses checked by him determines his barrier score on the test.

(e) *Tomkins-Horn picture arrangement test:* This test is designed for group administration. It consists of twenty-five plates, each containing three sketches that may be arranged in different ways to portray sequence of events. The respondent is asked to arrange them in a sequence which he considers as reasonable. The responses are interpreted as providing evidence confirming certain norms, respondent's attitudes, etc.

(vi) *Play techniques:* Under play techniques subjects are asked to improvise or act out a situation in which they have been assigned various roles. The researcher may observe such traits as hostility, dominance, sympathy, prejudice or the absence of such traits. These techniques have been used for knowing the attitudes of younger ones through manipulation of dolls. Dolls representing different racial groups are usually given to children who are allowed to play with them freely. The manner in which children organise dolls would indicate their attitude towards the class of persons represented by dolls. This is also known as *doll-play test*, and is used frequently in studies pertaining to sociology. The choice of colour, form, words, the sense of orderliness and other reactions may provide opportunities to infer deep-seated feelings.

(vii) *Quizzes, tests and examinations:* This is also a technique of extracting information regarding specific ability of candidates indirectly. In this procedure both long and short questions are framed to test through them the memorising and analytical ability of candidates.

(viii) *Sociometry:* Sociometry is a technique for describing the social relationships among individuals in a group. In an indirect way, sociometry attempts to describe attractions or repulsions between

[1] S.L. Dass, "*Personality Assessment Through Projective Movie Pictures*", p. 17.

individuals by asking them to indicate whom they would choose or reject in various situations. Thus, sociometry is a new technique of studying the underlying motives of respondents. "Under this an attempt is made to trace the flow of information amongst groups and then examine the ways in which new ideas are diffused. Sociograms are constructed to identify leaders and followers."[2] Sociograms are charts that depict the sociometric choices. There are many versions of the sociogram pattern and the reader is suggested to consult specialised references on sociometry for the purpose. This approach has been applied to the diffusion of ideas on drugs amongst medical practitioners.

7. Depth interviews: Depth interviews are those interviews that are designed to discover underlying motives and desires and are often used in motivational research. Such interviews are held to explore needs, desires and feelings of respondents. In other words, they aim to elicit unconscious as also other types of material relating especially to personality dynamics and motivations. As such, depth interviews require great skill on the part of the interviewer and at the same time involve considerable time. Unless the researcher has specialised training, depth interviewing should not be attempted.

Depth interview may be projective in nature or it may be a non-projective interview. The difference lies in the nature of the questions asked. Indirect questions on seemingly irrelevant subjects provide information that can be related to the informant's behaviour or attitude towards the subject under study. Thus, for instance, the informant may be asked on his frequency of air travel and he might again be asked at a later stage to narrate his opinion concerning the feelings of relatives of some other man who gets killed in an airplane accident. Reluctance to fly can then be related to replies to questions of the latter nature. If the depth interview involves questions of such type, the same may be treated as projective depth interview. But in order to be useful, depth interviews do not necessarily have to be projective in nature; even non-projective depth interviews can reveal important aspects of psycho-social situation for understanding the attitudes of people.

8. Content-analysis: Content-analysis consists of analysing the contents of documentary materials such as books, magazines, newspapers and the contents of all other verbal materials which can be either spoken or printed. Content-analysis prior to 1940's was mostly quantitative analysis of documentary materials concerning certain characteristics that can be identified and counted. But since 1950's content-analysis is mostly qualitative analysis concerning the general import or message of the existing documents. "The difference is somewhat like that between a casual interview and depth interviewing."[3] Bernard Berelson's name is often associated with. the latter type of content-analysis. "Content-analysis is measurement through proportion.... Content analysis measures pervasiveness and that is sometimes an index of the intensity of the force."[4]

The analysis of content is a central activity whenever one is concerned with the study of the nature of the verbal materials. A review of research in any area, for instance, involves the analysis of the contents of research articles that have been published. The analysis may be at a relatively simple level or may be a subtle one. It is at a simple level when we pursue it on the basis of certain characteristics of the document or verbal materials that can be identified and counted (such as on the basis of major scientific concepts in a book). It is at a subtle level when researcher makes a study of the attitude, say of the press towards education by feature writers.

[2] G.B. Giles, *Marketing*, p. 40–41.

[3] Carter V. Good and Douglas E. Scates, *Methods of Research*, p. 670.

[4] *Ibid*., p. 670.

COLLECTION OF SECONDARY DATA

Secondary data means data that are already available i.e., they refer to the data which have already been collected and analysed by someone else. When the researcher utilises secondary data, then he has to look into various sources from where he can obtain them. In this case he is certainly not confronted with the problems that are usually associated with the collection of original data. Secondary data may either be published data or unpublished data. Usually published data are available in: (a) various publications of the central, state are local governments; (b) various publications of foreign governments or of international bodies and their subsidiary organisations; (c) technical and trade journals; (d) books, magazines and newspapers; (e) reports and publications of various associations connected with business and industry, banks, stock exchanges, etc.; (f) reports prepared by research scholars, universities, economists, etc. in different fields; and (g) public records and statistics, historical documents, and other sources of published information. The sources of unpublished data are many; they may be found in diaries, letters, unpublished biographies and autobiographies and also may be available with scholars and research workers, trade associations, labour bureaus and other public/ private individuals and organisations.

Researcher must be very careful in using secondary data. He must make a minute scrutiny because it is just possible that the secondary data may be unsuitable or may be inadequate in the context of the problem which the researcher wants to study. In this connection Dr. A.L. Bowley very aptly observes that it is never safe to take published statistics at their face value without knowing their meaning and limitations and it is always necessary to criticise arguments that can be based on them.

By way of caution, the researcher, before using secondary data, must see that they possess following characteristics:

1. Reliability of data: The reliability can be tested by finding out such things about the said data: (a) Who collected the data? (b) What were the sources of data? (c) Were they collected by using proper methods (d) At what time were they collected?(e) Was there any bias of the compiler? (t) What level of accuracy was desired? Was it achieved ?

2. Suitability of data: The data that are suitable for one enquiry may not necessarily be found suitable in another enquiry. Hence, if the available data are found to be unsuitable, they should not be used by the researcher. In this context, the researcher must very carefully scrutinise the definition of various terms and units of collection used at the time of collecting the data from the primary source originally. Similarly, the object, scope and nature of the original enquiry must also be studied. If the researcher finds differences in these, the data will remain unsuitable for the present enquiry and should not be used.

3. Adequacy of data: If the level of accuracy achieved in data is found inadequate for the purpose of the present enquiry, they will be considered as inadequate and should not be used by the researcher. The data will also be considered inadequate, if they are related to an area which may be either narrower or wider than the area of the present enquiry.

From all this we can say that it is very risky to use the already available data. The already available data should be used by the researcher only when he finds them reliable, suitable and adequate. But he should not blindly discard the use of such data if they are readily available from authentic sources and are also suitable and adequate for in that case it will not be economical to

spend time and energy in field surveys for collecting information. At times, there may be wealth of usable information in the already available data which must be used by an intelligent researcher but with due precaution.

SELECTION OF APPROPRIATE METHOD FOR DATA COLLECTION

Thus, there are various methods of data collection. As such the researcher must judiciously select the method/methods for his own study, keeping in view the following factors:

1. **Nature, scope and object of enquiry:** This constitutes the most important factor affecting the choice of a particular method. The method selected should be such that it suits the type of enquiry that is to be conducted by the researcher. This factor is also important in deciding whether the data already available (secondary data) are to be used or the data not yet available (primary data) are to be collected.

2. **Availability of funds:** Availability of funds for the research project determines to a large extent the method to be used for the collection of data. When funds at the disposal of the researcher are very limited, he will have to select a comparatively cheaper method which may not be as efficient and effective as some other costly method. Finance, in fact, is a big constraint in practice and the researcher has to act within this limitation.

3. **Time factor:** Availability of time has also to be taken into account in deciding a particular method of data collection. Some methods take relatively more time, whereas with others the data can be collected in a comparatively shorter duration. The time at the disposal of the researcher, thus, affects the selection of the method by which the data are to be collected.

4. **Precision required:** Precision required is yet another important factor to be considered at the time of selecting the method of collection of data.

But one must always remember that each method of data collection has its uses and none is superior in all situations. For instance, telephone interview method may be considered appropriate (assuming telephone population) if funds are restricted, time is also restricted and the data is to be collected in respect of few items with or without a certain degree of precision. In case funds permit and more information is desired, personal interview method may be said to be relatively better. In case time is ample, funds are limited and much information is to be gathered with no precision, then mail-questionnaire method can be regarded more reasonable. When funds are ample, time is also ample and much information with no precision is to be collected, then either personal interview or the mail-questionnaire or the joint use of these two methods may be taken as an appropriate method of collecting data. Where a wide geographic area is to be covered, the use of mail-questionnaires supplemented by personal interviews will yield more reliable results per rupee spent than either method alone. The secondary data may be used in case the researcher finds them reliable, adequate and appropriate for his research. While studying motivating influences in market researches or studying people's attitudes in psychological/social surveys, we can resort to the use of one or more of the projective techniques stated earlier. Such techniques are of immense value in case the reason is obtainable from the respondent who knows the reason but does not want to admit it or the reason relates to some underlying psychological attitude and the respondent is not aware of it. But when the respondent knows the reason and can tell the same if asked, than a non projective questionnaire,

using direct questions, may yield satisfactory results even in case of attitude surveys. Since projective techniques are as yet in an early stage of development and with the validity of many of them remaining an open question, it is usually considered better to rely on the straight forward statistical methods with only supplementary use of projective techniques. Nevertheless, in pre-testing and in searching for hypotheses they can be highly valuable.

Thus, the most desirable approach with regard to the selection of the method depends on the nature of the particular problem and on the time and resources (money and personnel) available along with the desired degree of accuracy. But, over and above all this, much depends upon the ability and experience of the researcher. Dr. A.L. Bowley's remark in this context is very appropriate when he says that "in collection of statistical data common sense is the chief requisite and experience the chief teacher."

CASE STUDY METHOD

Meaning: The case study method is a very popular form of qualitative analysis and involves a careful and complete observation of a social unit, be that unit a person, a family, an institution, a cultural group or even the entire community. It is a method of study in depth rather than breadth. The case study places more emphasis on the full analysis of a limited number of events or conditions and their interrelations. The case study deals with the processes that take place and their interrelationship. Thus, case study is essentially an intensive investigation of the particular unit under consideration. The object of the case study method is to locate the factors that account for the behaviour-patterns of the given unit as an integrated totality.

According to H. Odum, "The case study method is a technique by which individual factor whether it be an institution or just an episode in the life of an individual or a group is analysed in its relationship to any other in the group."[5] Thus, a fairly exhaustive study of a person (as to what he does and has done, what he thinks he does and had done and what he expects to do and says he ought to do) or group is called a life or case history. Burgess has used the words "the social microscope" for the case study method."[6] Pauline V. Young describes case study as "a comprehensive study of a social unit be that unit a person, a group, a social institution, a district or a community."[7] In brief, we can say that case study method is a form of qualitative analysis where in careful and complete observation of an individual or a situation or an institution is done; efforts are made to study each and every aspect of the concerning unit in minute details and then from case data generalisations and inferences are drawn.

Characteristics: The important characteristics of the case study method are as under:

1. Under this method the researcher can take one single social unit or more of such units for his study purpose; he may even take a situation to study the same comprehensively.
2. Here the selected unit is studied intensively i.e., it is studied in minute details. Generally, the study extends over a long period of time to ascertain the natural history of the unit so as to obtain enough information for drawing correct inferences.

[5] H. Odum, *An Introduction to Social Research*, p. 229.

[6] Burgess, *Research Methods in Sociology*, p. 26 in Georges Gurvitch and W.E. Moore (Eds.) *Twentieth Century Sociology*.

[7] Pauline V. Young, *Scientific Social Surveys and Research*, p. 247.

3 In the context of this method we make complete study of the social unit covering all facets. Through this method we try to understand the complex of factors that are operative within a social unit as an integrated totality.

4 Under this method the approach happens to be qualitative and not quantitative. Mere quantitative information is not collected. Every possible effort is made to collect information concerning all aspects of life. As such, case study deepens our perception and gives us a clear insight into life. For instance, under this method we not only study how many crimes a man has done but shall peep into the factors that forced him to commit crimes when we are making a case study of a man as a criminal. The objective of the study may be to suggest ways to reform the criminal.

5 In respect of the case study method an effort is made to know the mutual inter-relationship of causal factors.

6 Under case study method the behaviour pattern of the concerning unit is studied directly and not by an indirect and abstract approach.

7 Case study method results in fruitful hypotheses along with the data which may be helpful in testing them, and thus it enables the generalised knowledge to get richer and richer. In its absence, generalised social science may get handicapped.

Evolution and scope: The case study method is a widely used systematic field research technique in sociology these days. The credit for introducing this method to the field of social investigation goes to Frederic Le Play who used it as a hand-maiden to statistics in his studies of family budgets. Herbert Spencer was the first to use case material in his comparative study of different cultures. Dr. William Healy resorted to this method in his study of juvenile delinquency, and considered it as a better method over and above the mere use of statistical data. Similarly, anthropologists, historians, novelists and dramatists have used this method concerning problems pertaining to their areas of interests. Even management experts use case study methods for getting clues to several management problems. In brief, case study method is being used in several disciplines. Not only this, its use is increasing day by day.

Assumptions: The case study method is based on several assumptions. The important assumptions may be listed as follows:

(i) The assumption of uniformity in the basic human nature in spite of the fact that human behaviour may vary according to situations.

(ii) The assumption of studying the natural history of the unit concerned.

(iii) The assumption of comprehensive study of the unit concerned.

Major phases involved: Major phases involved in case study are as follows:

(i) Recognition and determination of the status of the phenomenon to be investigated or the unit of attention.

(ii) Collection of data, examination and history of the given phenomenon.

(iii) Diagnosis and identification of causal factors as a basis for remedial or developmental treatment.

(iv) Application of remedial measures i.e., treatment and therapy (this phase is often characterised as case work).

(v) Follow-up programme to determine effectiveness of the treatment applied.

Advantages: There are several advantages of the case study method that follow from the various characteristics outlined above. Mention may be made here of the important advantages.

(i) Being an exhaustive study of a social unit, the case study method enables us to understand fully the behaviour pattern of the concerned unit. In the words of Charles Horton Cooley, "case study deepens our perception and gives us a clearer insight into life.... It gets at behaviour directly and not by an indirect and abstract approach."

(ii) Through case study a researcher can obtain a real and enlightened record of personal experiences which would reveal man's inner strivings, tensions and motivations that drive him to action along with the forces that direct him to adopt a certain pattern of behaviour.

(iii) This method enables the researcher to trace out the natural history of the social unit and its relationship with the social factors and the forces involved in its surrounding environment.

(iv) It helps in formulating relevant hypotheses along with the data which may be helpful in testing them. Case studies, thus, enable the generalised knowledge to get richer and richer.

(v) The method facilitates intensive study of social units which is generally not possible if we use either the observation method or the method of collecting information through schedules. This is the reason why case study method is being frequently used, particularly in social researches.

(vi) Information collected under the case study method helps a lot to the researcher in the task of constructing the appropriate questionnaire or schedule for the said task requires thorough knowledge of the concerning universe.

(vii) The researcher can use one or more of the several research methods under the case study method depending upon the prevalent circumstances. In other words, the use of different methods such as depth interviews, questionnaires, documents, study reports of individuals, letters, and the like is possible under case study method.

(viii) Case study method has proved beneficial in determining the nature of units to be studied along with the nature of the universe. This is the reason why at times the case study method is alternatively known as "mode of organising data".

(ix) This method is a means to well understand the past of a social unit because of its emphasis of historical analysis. Besides, it is also a technique to suggest measures for improvement in the context of the present environment of the concerned social units.

(x) Case studies constitute the perfect type of sociological material as they represent a real record of personal experiences which very often escape the attention of most of the skilled researchers using other techniques.

(xi) Case study method enhances the experience of the researcher and this in turn increases his analysing ability and skill.

(xii) This method makes possible the study of social changes. On account of the minute study of the different facets of a social unit, the researcher can well understand the social change then and now. This also facilitates the drawing of inferences and helps in maintaining the continuity of the research process. In fact, it may be considered the gateway to and at the same time the final destination of abstract knowledge.

(xiii) Case study techniques are indispensable for therapeutic and administrative purposes. They are also of immense value in taking decisions regarding several management problems. Case data are quite useful for diagnosis, therapy and other practical case problems.

Limitations: Important limitations of the case study method may as well be highlighted.

(i) Case situations are seldom comparable and as such the information gathered in case studies is often not comparable. Since the subject under case study tells history in his own words, logical concepts and units of scientific classification have to be read into it or out of it by the investigator.

(ii) Read Bain does not consider the case data as significant scientific data since they do not provide knowledge of the "impersonal, universal, non-ethical, non-practical, repetitive aspects of phenomena."[8] Real information is often not collected because the subjectivity of the researcher does enter in the collection of information in a case study.

(iii) The danger of false generalisation is always there in view of the fact that no set rules are followed in collection of the information and only few units are studied.

(iv) It consumes more time and requires lot of expenditure. More time is needed under case study method since one studies the natural history cycles of social units and that too minutely.

(v) The case data are often vitiated because the subject, according to Read Bain, may write what he thinks the investigator wants; and the greater the rapport, the more subjective the whole process is.

(vi) Case study method is based on several assumptions which may not be very realistic at times, and as such the usefulness of case data is always subject to doubt.

(vii) Case study method can be used only in a limited sphere., it is not possible to use it in case of a big society. Sampling is also not possible under a case study method.

(viii) Response of the investigator is an important limitation of the case study method. He often thinks that he has full knowledge of the unit and can himself answer about it. In case the same is not true, then consequences follow. In fact, this is more the fault of the researcher rather than that of the case method.

Conclusion: Despite the above stated limitations, we find that case studies are being undertaken in several disciplines, particularly in sociology, as a tool of scientific research in view of the several advantages indicated earlier. Most of the limitations can be removed if researchers are always conscious of these and are well trained in the modern methods of collecting case data and in the scientific techniques of assembling, classifying and processing the same. Besides, case studies, in modern times, can be conducted in such a manner that the data are amenable to quantification and statistical treatment. Possibly, this is also the reason why case studies are becoming popular day by day.

Question

1. Enumerate the different methods of collecting data. Which one is the most suitable for conducting enquiry regarding family welfare programme in India? Explain its merits and demerits.

[8] Pauline V. Young, *Scientific social surveys and research*, p. 262.

2. "It is never safe to take published statistics at their face value without knowing their meaning and limitations." Elucidate this statement by enumerating and explaining the various points which you would consider before using any published data. Illustrate your answer by examples wherever possible.
3. Examine the merits and limitations of the observation method in collecting material. Illustrate your answer with suitable examples.
4. Describe some of the major projective techniques and evaluate their significance as tools of scientific social research.
5. How does the case study method differ from the survey method? Analyse the merits and limitations of case study method in sociological research.
6. Clearly explain the difference between collection of data through questionnaires and schedules.
7. Discuss interview as a technique of data collection.
8. Write short notes on:
   (a) Depth interviews;
   (b) Important aspects of a questionnaire;
   (c) Pantry and store audits;
   (d) Thematic Apperception Test;
   (e) Holtzman Inkbolt Test.
9. What are the guiding considerations in the construction of questionnaire? Explain.
10. Critically examine the following:
   (i) Interviews introduce more bias than does the use of questionnaire.
   (ii) Data collection through projective techniques is considered relatively more reliable.
   (iii) In collection of statistical data commonsense is the chief requisite and experience the chief teacher.
11. Distinguish between an experiment and survey. Explain fully the survey method of research.

[M. Phi. (EAFM) Exam. 1987 Raj. Uni.]

12. "Experimental method of research is not suitable in management field." Discuss, what are the problems in the introduction of this research design in business organisation?

[M.B.A. (Part I) Exam. 1985 Raj. Uni.]

Appendix (i)

Guidelines for Constructing Questionnaire/Schedule

The researcher must pay attention to the following points in constructing an appropriate and effective questionnaire or a schedule:

1. The researcher must keep in view the problem he is to study for it provides the starting point for developing the Questionnaire/Schedule. He must be clear about the various aspects of his research problem to be dealt with in the course of his research project.
2. Appropriate form of questions depends on the nature of information sought, the sampled respondents and the kind of analysis intended. The researcher must decide whether to use closed or open-ended question. Questions should be simple and must be constructed with a view to their forming a logical part of a well thought out tabulation plan. The units of enumeration should also be defined precisely so that they can ensure accurate and full information.
3. Rough draft of the Questionnaire/Schedule be prepared, giving due thought to the appropriate sequence of putting questions. Questionnaires or schedules previously drafted (if available) may as well be looked into at this stage.
4. Researcher must invariably re-examine, and in case of need may revise the rough draft for a better one. Technical defects must be minutely scrutinised and removed.
5. Pilot study should be undertaken for pre-testing the questionnaire. The questionnaire may be edited in the light of the results of the pilot study.
6. Questionnaire must contain simple but straight forward directions for the respondents so that they may not feel any difficulty in answering the questions.

Appendix (ii)

Guidelines for Successful Interviewing

Interviewing is an art and one learns it by experience. However, the following points may be kept in view by an interviewer for eliciting the desired information:

1. Interviewer must plan in advance and should fully know the problem under consideration. He must choose a suitable time and place so that the interviewee may be at ease during the interview period. For this purpose some knowledge of the daily routine of the interviewee is essential.
2. Interviewer's approach must be friendly and informal. Initially friendly greetings in accordance with the cultural pattern of the interviewee should be exchanged and then the purpose of the interview should be explained.
3. All possible effort should be made to establish proper rapport with the interviewee; people are motivated to communicate when the atmosphere is favourable.
4. Interviewer must know that ability to listen with understanding, respect and curiosity is the gateway to communication, and hence must act accordingly during the interview. For all this, the interviewer must be intelligent and must be a man with self-restraint and self-discipline.
5. To the extent possible there should be a free-flowing interview and the questions must be well phrased in order to have full cooperation of the interviewee. But the interviewer must control the course of the interview in accordance with the objective of the study.
6. In case of big enquiries, where the task of collecting information is to be accomplished by several interviewers, there should be an interview guide to be observed by all so as to ensure reasonable uniformity in respect of all salient points in the study.

Appendix (iii)

Difference Between Survey and Experiment

The following points are noteworthy so far as difference between survey and experiment is concerned:

(i) Surveys are conducted in case of descriptive research studies where as experiments are a part of experimental research studies.

(ii) Survey-type research studies usually have larger samples because the percentage of responses generally happens to be low, as low as 20 to 30%, especially in mailed questionnaire studies. Thus, the survey method gathers data from a relatively large number of cases at a particular time; it is essentially cross-sectional. As against this, experimental studies generally need small samples.

(iii) Surveys are concerned with describing, recording, analysing and interpreting conditions that either exist or existed. The researcher does not manipulate the variable or arrange for events to happen. Surveys are only concerned with conditions or relationships that exist, opinions that are held, processes that are going on, effects that are evident or trends that are developing. They are primarily concerned with the present but at times do consider past events and influences as they relate to current conditions. Thus, in surveys, variables that exist or have already occurred are selected and observed.

Experimental research provides a systematic and logical method for answering the question, "What will happen if this is done when certain variables are carefully controlled or manipulated?" In fact, deliberate manipulation is a part of the experimental method. In an experiment, the researcher measures the effects of an experiment which he conducts intentionally.

(iv) Surveys are usually appropriate in case of social and behavioural sciences (because many types of behaviour that interest the researcher cannot be arranged in a realistic setting) where as experiments are mostly an essential feature of physical and natural sciences.

(v) Surveys are an example of field research where as experiments generally constitute an example of laboratory research.

(vi) Surveys are concerned with hypothesis formulation and testing the analysis of the relationship between non-manipulated variables. Experimentation provides a method of hypothesis testing. After experimenters define a problem, they propose a hypothesis. They then test the hypothesis and confirm or disconfirm it in the light of the controlled variable relationship that they have observed. The confirmation or rejection is always stated in terms of probability rather than certainty. Experimentation, thus, is the most sophisticated, exacting and powerful method for discovering and developing an organised body of knowledge. The ultimate purpose of experimentation is to generalise the variable relationships so that they may be applied outside the laboratory to a wider population of interest.*

(vii) Surveys may either be census or sample surveys. They may also be classified as social surveys, economic surveys or public opinion surveys. Whatever be their type, the method of data collection happens to be either observation, or interview or questionnaire/opinionnaire or some projective technique(s). Case study method can as well be used. But in case of experiments, data are collected from several readings of experiments.

(viii) In case of surveys, research design must be rigid, must make enough provision for protection against bias and must maximise reliability as the aim happens to be to obtain complete and accurate information. Research design in case of experimental studies, apart reducing bias and ensuring reliability, must permit drawing inferences about causality.

(ix) Possible relationships between the data and the unknowns in the universe can be studied through surveys where as experiments are meant to determine such relationships.

(x) Causal analysis is considered relatively more important in experiments where as in most social and business surveys our interest lies in understanding and controlling relationships between variables and as such correlation analysis is relatively more important in surveys.

*John W. Best and James V. Kahn, "*Research in Education*", 5th ed., Prentice-Hall of India Pvt. Ltd., New Delhi, 1986, p.111.